Continual Red Queen Evolution in Nonreciprocal Ecological Models

Aditya Mahadevan

Department of Physics, Stanford University, Stanford, CA 94305, USA

Daniel S. Fisher

Department of Applied Physics, Stanford University, Stanford, CA 94305, USA

(Dated: October 31, 2024)

Feedbacks between evolution and ecology are ubiquitous, with ecological interactions determining which mutants are successful, and these mutants in turn modifying community structure. We study the evolutionary dynamics of several ecological models with overlapping niches, including consumer resource and Lotka-Volterra models. Evolution is assumed slow, with ecological dynamics producing a stable community (with some strains going permanently extinct) between successive introductions of an invader or mutant. When new strains are slowly added to the community, either from an external pool or by large-effect mutations, the ecosystem converges, after an initial evolutionary transient, to a diverse eco-evolutionary steady state. In this "Red Queen" phase, of continual evolution, the biodiversity continues to turn over without the invasion probability of new variants getting any smaller. For resource-mediated interactions, the Red Queen phase obtains for any amount of asymmetry in the interactions between strains, and it is robust to "general fitness" differences in the intrinsic growth rates of strains. Via a dynamical mean field theory framework valid for high-dimensional phenotype space, we analytically characterize the Red Queen eco-evolutionary steady state for a particular limit of model parameters. Scaling arguments enable a more general understanding of the steady state and evolutionary transients toward it. This work therefore establishes simple models of continual evolution in an ecological context *without* host-pathogen arms races, and points to the generality of Red Queen evolution. However other phases are also found in simple models: For generalized Lotka-Volterra models with weakly asymmetric interactions an "oligarch" phase emerges in which the evolutionary dynamics continually slow down and a substantial fraction of the community's abundance condenses into a handful of slowly turning-over strains.

CONTENTS

I. INTRODUCTION

Ecological and evolutionary processes are inextricably linked. In diverse microbial communities with large population sizes, there is constant interplay between ecological interactions and new genetic variants. Evolution continually creates new variants that interact with the community differently from their parent strains, and these ecological interactions determine the community's properties and the fate of the new mutants, thereby influencing future evolution. As the increasing body of data from natural microbial communities shows, microbes live in close proximity with

one another, with multiple strains and substrains of the same species coexisting in many natural environments [1–3]. Understanding how ecological interaction webs influence evolution—and vice versa—is crucial for developing a richer picture of microbial evolution in both experimental and natural settings. A major puzzle is the ubiquity of closely related microbial strains coexisting, as evinced by the broad distribution of genomic divergences found within bacterial species apparently occupying the same ecological niche [4, 5]. How does such diversity evolve and persist without clear ecological differences stabilizing the coexistence of strains? Theory is needed to develop distinguishable scenarios for the evolution and maintenance of fine-scale biodiversity in large microbial populations.

Even in controlled laboratory experiments with initially-clonal microbial populations, the development of ecology is inevitable. Although *apparent* selection pressures in the laboratory can be held constant, evolution in "simple" environments still generically results in the diversification and coexistence of different strains [6–8]. Long after increases in fitness—as measured by competition with the ancestor—have slowed, and the population seems to be well adapted to laboratory conditions, it has been observed in the *E. coli* experiments of Richard Lenski, that evolution continues without further decrease in the rate of successful mutations [9]. Diversification and coexistence of lineages is observed even in such experiments *designed* to be simple enough to eliminate ecological effects [9, 10]. Remarkably, even in laboratory experiments on bacterial populations without an external nutrient supply—which might be expected to exhibit severe resource limitation and diversity collapse—diversification and phenotypic turnover is found to continue over a timescale of years [11].

With interacting bacteria and phage populations, there is clear evidence—both experimental and observational—for the continual turnover and diversification of microbial lineages [12–14]. The "Red Queen hypothesis" [15] is often invoked to describe such arms-race evolutionary dynamics of hosts and pathogens. In the current work we use "Red Queen dynamics" more generally to describe continual turnover of biodiversity in a community without successful invasions of new mutants becoming rarer, and we look for such dynamics in other ecological contexts.

Inspired by observations of microbial populations, we aim to understand what generic scenarios can emerge in the long term dynamics of evolution of ecological communities. We are particularly interested in robust behaviors that obtain across a range of parameters and models—such "phases" of eco-evolutionary dynamics present a set of qualitative possibilities that one can hope to compare with experiment or observation. The phases that we study emerge in the limit of large biological complexity: *high-dimensional phenotype and interaction spaces* upon which evolutionary and ecological processes play out. Looking for phases that are robust in the limit of high phenotypic dimensionality D—roughly the number of phenotypic properties of organisms that affect the ecological dynamics—increases the likelihood that the conclusions we draw are somewhat universal and relevant to natural systems, which are more complex—in multiple ways—than the simple models we study.

In order to understand what long term behaviors emerge in high-dimensional eco-evolutionary models, it is useful to first think separately about the dynamics of evolution and ecology, initially in a well-mixed population with *no spatial structure*. A further simplification is natural for large microbial populations: that the dynamics are essentially deterministic, with demographic fluctuations negligible except close to extinctions.

Evolution without ecology: In a static environment without ecology, gradual evolution of a large almost-clonal population can be caricatured as gradient ascent on a complex fitness landscape. The fitness landscape is a map from some high-dimensional organismal phenotype to the "fitness": a misleading but canonized term for a measure of reproductive success that is surely not a single quantity except in the simplest non-ecological contexts. The emergence and fixation of beneficial mutations drives the population to high-fitness regions of phenotype space and evolution slows down as the population approaches one of the (typically) many local maxima of the fitness landscape [16]. However in high dimensions, this simple intuition is only part of the story: the phenotype will generally spend a long time wandering around saddles before "committing" to a particular maximum, which is likely to be only marginally stable [17].

In large populations with an abundant supply of mutations, the evolutionary dynamics of the population of genotypes is complex, with simultaneous exploration of multiple phenotypic directions: this regime has not been much studied theoretically. However if there is no ecology except resource limitation capping the total population size, eventually the supply of beneficial mutations will become so limited that the population will again become almost clonal. By contrast, in an extrinsically changing environment, the fitness landscape depends on time, and is better caricatured as a "seascape." Not surprisingly, a continually changing environment can lead to evolutionary dynamics which proceed indefinitely without slowing down [18] and produce continually turning-over diversity: again, such phenomena have not been much studied for large populations in high phenotypic dimension.

Ecology without evolution: Just as evolutionary dynamics of large populations can display rich behavior even without ecology, ecological dynamics with no evolution can exhibit various behaviors, with globally stable, multistable, and chaotic states all possible in simple models [19–21]. The most-studied contexts are ecological dynamics of *top-down assembled* communities, which consist of a large number of species brought together with their abundances obeying some prescribed ecological dynamics. The simplest ecological phase that obtains at long times is a state in which species abundances reach a globally attracting stable fixed point with some number of species (the stable community) having nonzero abundance and the rest going extinct, unable to reinvade even if reintroduced. As the diversity of this assembled community increases, the stable phase can yield to a more complicated behavior exhibiting multiple attractors and transient chaos, which remains only partially understood [19, 21, 22]. Chaotic fluctuations in the high-diversity unstable regime lead to many extinctions and the system settles down to one of many possible stable communities or limit cycles, but these fixed points are generically unstable to invasion by some of the extinct strains. A pop-

ulation of dormant spores that germinate and resuscitate extinct populations can sustain the chaos for times exponentially long in the lifetimes of the spores, but as a matter of principle such dynamics do not constitute a true ecological phase [23].

Interplay of ecology and evolution: To bring evolution and ecology together, we focus here on the slowly evolving limit in which the ecological dynamics are much faster than the evolutionary dynamics and spatial mixing is fast enough to be ignored. Motivated especially by the puzzle of fine-scale diversity, we are primarily interested in multiple strains of a single species with overlapping niches. This regime is the hardest in which to find sustained diversity and continual evolution because it does not allow short-term transient coexistence driven by continual invasions of new mutants, which was studied in refs. [24–26].

We study models in which the ecological dynamics are simplest, converging to a (usually) globally attracting fixed point with some strains permanently extinct. Under slow evolution, do the resulting eco-evolutionary dynamics remain diverse and continually turn over in a "Red Queen" steady state, or does evolution get progressively slower and slower? Do generalist mutants take over and cause the diversity to crash? Are there multiple different phases of eco-evolutionary dynamics for different parameter regimes?

It has been shown that Red Queen dynamics emerge generically in models of an almost-clonal population evolving in a high-dimensional fitness landscape with a small amount of environmental feedback [17]. In these models, the fitness landscape—now better caricatured as a "snowscape"—changes as the evolving population affects its environment. For even a small amount of such feedback, the population wanders through phenotype space without settling down to a stable fixed point, due to selective forces that arise from the gradient of its continually-modified fitness landscape. Universality emerges in these models in the limit of large phenotype dimension, with different microscopic models displaying similar phase diagrams, including a phase of sustained evolutionary chaos [17].

In the eco-evolutionary models we study here, ecology acts loosely as a manifestation of feedback between organism and environment, with the "environment" of any particular strain set by its interactions with the rest of the community, and thus subject to change every time a new mutant arises in the community. Can natural models of this process result in a Red Queen phase of evolution? If so, what sets the diversity of the community in this phase?

Some previous work has found that ecological communities, although capable of sustaining extensive diversity through standard mechanisms of resource competition [27] lose much of their diversity when they start to gradually evolve either by small mutations or by the introduction of independent species [28, 29]. These results were argued to deepen the puzzle of the extensive diversity observed in nature. However, the models in which such results have been reported are complicated enough that analytical understanding is difficult, and generality of the claims is not clear—indeed other work has suggested that phenotypic complexity *allows* the gradual evolution of high diversity [30].

Evolutionary dynamics: The evolutionary process we study involves adding strains one by one to an ecological community, with extinctions permanent when they occur. Once extinct, a strain will never be able to re-emerge, even if conditions become favorable for its growth. Therefore the order in which strains are introduced influences the fate of the community. However, the *statistical* properties of the evolving community for large phenotypic dimension D depend only on the pool from which invaders are drawn, rather than on the specific properties of the invaders.

Previous work has studied the process of adding strains into ecological models *without* permanent extinctions [31–33], where re-invasion of previously-extinct strains is possible if conditions later become favorable for them. This process eliminates the important difference between communities assembled all at once (top-down) and gradually (bottom-up): the latter is our focus.

A large amount of work by theoretical ecologists and statistical physicists has focused on ecological models that are nongeneric because they admit a Lyapunov function which is always increasing under the eco-evolutionary dynamics [34–37]. The existence of such a Lyapunov function depends on the symmetry of ecological interactions [38], e.g. through perfect efficiency in the conversion of nutrients to biomass, which does not hold in biological systems. Therefore the generality of conclusions drawn from such models is unclear. In the current work we focus on generic behavior, drawing conclusions away from the special slice of parameter space with a Lyapunov function. Indeed we will show that, under evolution, the behavior of models with a Lyapunov function is qualitatively different from the general case. Several other studies [39–41] have noted the importance of *nonreciprocal* or *asymmetric* ecological interactions for reaching a state of continual evolution—we will provide further evidence for this conclusion in the high-diversity limit, which complements the more often-studied low-diversity case [39, 40]

Models: We are interested in dynamics of closely related strains, where differences between strain interactions with one another are sums of many positive and negative effects. To caricature this biological complexity, we parameterize the strain-to-strain differences in phenotypic properties as *random variables* drawn from some distribution. This approximation of unknown complexity by randomness is a useful way to get a sense of general patterns of behavior that might apply across a range of ecological-evolutionary situations [42].

In the simplest models, strains interact with one another *directly*, with all pairwise interactions chosen from some specified ensemble. These models, which include generalized Lotka-Volterra interactions [19], suffer from a lack of strain phenotypes which, when present, dictate how each strain interacts with any other prospective strain. As a result, the effective phenotypic dimensionality in Lotka-Volterra models is infinite. Problematically, such models do not admit a straightforward way of drawing random mutants that are correlated with a parent strain already in the community [43]. Nonetheless we will study the Lotka-Volterra eco-evolutionary dynamics in the case of independent invaders added to the community. This simple limit is instructive for understanding evolution driven by mutant strains that are correlated with their parents.

To explicitly study the effects of parent-mutant correlations we introduce structure into the ecological interactions, specifically considering the case where they are mediated by a collection of chemicals (or resources) in the environment. Each strain has a phenotype which specifies its effect on the resources and the resources' effect on its growth rate—these phenotypes consequently determine the inter-strain interactions. Small mutations correspond to having highly correlated parent and mutant phenotypes, with the similarity between mutant and parent in the community playing an important role.

The phenotype of each strain in such *consumer resource* models is given by its growth and consumption rates for each of the resources. The number of resources—or more generally, interaction-mediating chemicals—is the key quantity, D. This number sets the *dimensionality* of the environment and of the interactions between consumers. It also sets a natural scale for the number of strains that can coexist: in the models we study here, the diversity is bounded above by D, and "high-diversity" means that the number of coexisting strains is $\mathcal{O}(D)$. The limit of large D—surely biologically relevant [30]—will allow us to analyze statistical (hopefully universal) properties of the resulting eco-evolutionary dynamics.

When mutational effects are sufficiently large, we find that key aspects of the evolutionary dynamics are captured by the dynamics of *bottom-up assembly* (with extinctions), wherein phenotypically *independent* invaders enter the community serially. The understanding from this simplified regime gives us qualitative insight into differences between a top-down-assembled and an evolved community, along with a framework for understanding the Red Queen phase that occurs for correlated mutants.

In order for the Red Queen phase to be robust, it must be stable to the presence of "general fitness" mutations which increase the growth rate of a strain *regardless* of the composition of the community. These general fitnesses appear naturally in models of resource competition, where they correspond to the growth rates of strains in isolation. Metabolic tradeoffs, which constrain the width of the general fitness distribution, are sometimes invoked to sustain high levels of diversity in ecological settings. These tradeoffs can be strict [44, 45] or—more realistically—allow some variation in the total metabolic budget [36]. Without strict tradeoffs, can consumer resource and Lotka-Volterra models evolve towards a diverse state in which the distribution of extant general fitnesses is narrow? Or does evolution constantly push into the tail of the general fitness distribution and fail to reach a steady state? Do generalist species emerge and reduce diversity by outcompeting other strains?

Our main result is that there exists a robust high-diversity Red Queen phase of eco-evolutionary dynamics characterized by continual evolution and turnover of biodiversity, and that this phase occurs across a range of parameters and models. Importantly, the Red Queen phase persists in the presence of general fitness differences, as long as their distribution decays fast enough. The tail of the general fitness distribution controls the probability of successful invasion and therefore the rate of turnover in the steady state, but otherwise does not limit the diversity.

II. CONSUMER RESOURCE MODEL

A. Ecological dynamics and stable communities

We first study a standard model of resource competition with D externally-supplied resources (or more generally, interaction-mediating chemicals), and an assembled population of S different strains that consume these resources (see refs. [35, 36, 38, 44]). The ecological dynamics of this consumer resource (CR) model are

$$\frac{dn_i}{dt} = n_i \left(\sum_{\beta=1}^{D} \mathcal{G}_{i\beta} \mathcal{R}_\beta - d \right) \tag{1a}$$

$$\frac{d\mathcal{R}_\alpha}{dt} = K_\alpha - \mathcal{R}_\alpha \sum_{j=1}^{S} \mathcal{F}_{j\alpha} n_j - \omega \mathcal{R}_\alpha, \tag{1b}$$

where the $\{n_i\}$ are the populations of strains, labeled by Roman indices $i \in \{1, 2, \ldots S\}$ and the $\{\mathcal{R}_\alpha\}$ are resource abundances labeled by Greek indices $\alpha \in \{1, 2, \ldots D\}$. The resource supply vector $\mathbf{K}$ has components K_α, and d and ω are respectively the death rate of the consumers and the decay (or dilution) rate of the resources. Without substantial loss of generality, we take all the K_α to be equal and given by a constant K, and we set $\omega = 0$ to reduce the number of parameters without changing the qualitative behavior. The $S \times D$ matrices $\mathcal{G}$ and $\mathcal{F}$ respectively parameterize the *growth* and *feeding* rates of each consumer on each resource: the phenotype of strain i is the corresponding row of $\mathcal{G}$ and $\mathcal{F}$ taken together, so that the dimensionality of phenotype space is $2D$. The elements of $\mathcal{G}$ and $\mathcal{F}$ consist of some positive mean—an average growth or feeding rate—in addition to a heterogeneous part that varies between strains and resources. Ref. [36] studied the ecological dynamics of a *symmetric* version of this model with $\mathcal{F} = \mathcal{G}$. Although we expect $\mathcal{G}$ and $\mathcal{F}$ to be positively correlated, some amount of asymmetry—arising from, e.g., varying efficiencies in utilization of resources—is generic, and we will show that this asymmetry is important for the dynamics of interest to us.

For closely related strains, the differences between interactions will be sums of positive and negative effects: thus we treat $\mathcal{G}_{j\alpha}$ and $\mathcal{F}_{j\alpha}$ as random variables drawn from some distribution. With the interpretation of consumption and growth, one expects the elements of $\mathcal{F}$ and $\mathcal{G}$ to be positive, though a small fraction being negative (corresponding loosely to crossfeeding or inhibitory effects) does not change the behavior. For simplicity we draw the elements from gaussian distributions with means μ_g, μ_f and variances σ_g^2, σ_f^2 respectively. The choice of a gaussian distribution is convenient as we will later take advantage of this structure to evolve strains by drawing from the ensemble of interactions conditional on inequality constraints. However, as we show in Section IV E, the upper tails of the distributions of the elements of $\mathcal{G}$ and $\mathcal{F}$ play an important role in the evolution.

We parameterize elementwise correlations between the gaussian entries of $\mathcal{G}$ and $\mathcal{F}$ by $\mathrm{Corr}[\mathcal{G}_{j\beta}, \mathcal{F}_{j\alpha}] = \kappa \delta_{jk} \delta_{\alpha\beta}$, where $\mathrm{Corr}[A, B]$ is the Pearson correlation coefficient between A and B, and the *symmetry parameter* $\kappa \in [0, 1]$ parameterizes the similarity between growth and consumption rates. The ratios σ_g/μ_g and σ_f/μ_f determine what

fraction of the entries of $\mathcal{G}$ and $\mathcal{F}$ are positive; we consider $\sigma_f \lesssim \mu_f$ and $\sigma_g \lesssim \mu_g$ so that most entries are positive. With the choice of $\omega = 0$ there are only three important parameters other than S and D: κ, σ_g/μ_g and σ_f/μ_f, with K setting a scale for the consumer abundances (Appendix A). Without loss of generality we will set $\mu_f = \mu_g = 1$, $d = 1$ and $K = 1$.

The degree of correlation, κ, between the feeding and growth rates plays an important role: $\kappa = 1$ presumes identical efficiencies of conversion between resources and biomass between strains [36]. This symmetry guarantees the existence of a Lyapunov function (see Appendix A) that is always increasing along the ecological (and later evolutionary) dynamics of the CR model [38]. However for generic $\kappa < 1$ the dynamics do not have a Lyapunov function, which makes a big difference for the evolutionary behavior, as we will see shortly.

One must supplement the deterministic dynamics of Equation 1 with a small abundance threshold below which strains go *extinct*: their abundances are then set to zero and *cannot recover*. In large populations the dynamics and extinctions are driven deterministically by growth rate differences, so we neglect demographic noise. Starting from positive initial conditions, the strain abundances in the CR model generally reach a fixed point with some fraction of strains going extinct and others persisting at nonzero abundance. We denote by $\mathcal{C}$ the *community of extant strains*: those with positive fixed point abundances. For small enough S—less than some multiple of D—the fixed point is generally stable and cannot be invaded by any of the extinct strains. For large D, in the parameter ranges of interest, the stable uninvadable community is highly likely to be unique [38]. The breakdown of this uniqueness above a critical value of S (discussed for a related model in ref. [21]), will not play a role in the evolving communities of interest here.

At the ecological fixed point of the CR dynamics, the abundances $\{n_i^*\}$ of the extant strains and the resources $\{\mathcal{R}_\alpha^*\}$ are determined by

$$\mathcal{R}_\alpha^* = \frac{K_\alpha}{\sum_{j \in \mathcal{C}} \mathcal{F}_{j\alpha} n_j^*} \tag{2a}$$

$$0 = \sum_{\alpha=1}^{D} \mathcal{G}_{i\alpha} \mathcal{R}_\alpha^* - d \quad \text{for} \quad i \in \mathcal{C} \tag{2b}$$

$$n_j^* = 0 \text{ for } j \notin \mathcal{C} \tag{2c}$$

with the strains $j \notin \mathcal{C}$ extinct and unable to reinvade. Whether a strain can invade from low abundance is determined by the sign of its *invasion fitness*

$$\xi_i = \sum_{\alpha=1}^{D} \mathcal{G}_{i\alpha} \mathcal{R}_\alpha^* - d. \tag{3}$$

If ξ_i is negative, strain i cannot invade.

The total population size, $N \equiv \sum_i n_i^*$, is of order $\mu_g \sum_\alpha K_\alpha/\mu_f$, which is roughly D for our choice of parameters. (We have kept the dependence on $\mathbf{K}$ explicit because it carries a physical meaning and characterizes the environment. In our simulations all of its components are 1: the effects of changing $\mathbf{K}$ are discussed briefly in Section V.)

We define the *extant strain diversity*, $L = |\mathcal{C}|$, as the number of strains *left* in the community at the ecologically stable fixed point. In accordance with the competitive exclusion principle [27], the maximum L that can stably coexist at a fixed point is D, as seen from the insufficiency of the fixed point conditions otherwise. For natural choices—including ours—of the ensemble of interactions, the typical number of strains that can stably coexist at a fixed point scales linearly with D with a coefficient less than 1. This linear relationship between the diversity of the community and the number of resource-mediating chemicals is expected to be generic across a range of consumer resource models with heterogeneous interactions, with the statistics of the interactions affecting the prefactor in the scaling [36]. In ecological language, the number of *potential niches* is D but not all of these can be filled.

Measures of diversity: In the ecological literature many different measures of diversity of communities have been introduced [46]. We will use L, the total number of species in the community, which is sometimes called the *richness*. Other popular measures are the Shannon diversity: the weighted average of $\log(N/n_i)$, and the *inverse Simpson index* (*inverse participation ratio* in physics terminology), given by $N^2/(\sum_i n_i^2)$. If there is a broad distribution of abundances in a community, these measures can be quite different. However in our models, as we shall see, the distribution of abundances is not very broad so which measure we use does not affect our conclusions. For simplicity we will refer to L as the "diversity."

B. Evolutionary processes

Evolution by independent invaders: We now focus on slow evolution of stable communities in the CR model, assuming a clean separation of timescales between evolution and ecology. If the resource dynamics are fast enough, they can be integrated out to yield equations for the strain dynamics which reach the same fixed point irrespective of the resource timescale—we can therefore assume fast resource dynamics to simplify our dynamical equations without affecting the evolutionary behavior. The ecological dynamics of the community run until a fixed point is reached. Then a single new strain is introduced, and the ecological dynamics again run until a fixed point is reached. The ecological time scale thus does not affect the evolutionary dynamics of primary interest.

We first study the case of *independent invaders* into the community. Although this process is sometimes referred to as "bottom-up assembly," it will give us insight into evolutionary dynamics with large-effect mutations. We start with a single strain of consumer and a collection of D resources, and repeatedly add an unrelated consumer strain drawn from the gaussian phenotype ensemble. We define an *epoch* as the unit of time corresponding to a *successful* invasion followed by the relaxation of the ecological dynamics to a fixed point. Although the length of an epoch in ecological time varies between epochs, in our case of slow evolution the ecological timescale is not important.

With each new strain added, we find a stable uninvadable ecological fixed point of the community either by integrat-

ing the dynamical equations with a small extinction threshold, or by the more efficient iterative algorithm described in Appendix D. Once L is sufficiently large, we expect that, with high probability, this fixed point is the *unique* stable uninvadable fixed point for the collection of extant strains in each epoch—this intuition is based on analogy with the assembled community in which the uninvadable fixed point is unique as long S and L are of similar size [19], as is the case in our evolving system where the size of the extant community *plus* the new invader is analogous to S.

Invaders: In order for a new strain, labelled I, to be able to invade, it must have positive invasion fitness ξ_I. Furthermore, if it successfully invades, its fixed point abundance n_I^* will be approximately proportional to ξ_I, with the constant of proportionality determined by the properties of the community into which I invades, as can be seen by a perturbative analysis for large L. The dot product between the growth vector $\boldsymbol{\mathcal{G}}_I$ (the Ith row of $\mathcal{G}$) and the resource vector $\boldsymbol{\mathcal{R}}^*$ of the current community determines the invasion fitness of strain I. Therefore when we draw the phenotype of I—consisting of D-dimensional vectors $\boldsymbol{\mathcal{G}}_I$ and $\boldsymbol{\mathcal{F}}_I$—from their gaussian ensemble, we condition on $\xi_I > 0$, which drastically speeds up numerics when the evolution pushes the extant distribution of interactions to have large mean, making invaders with $\xi_I > 0$ very rare [47].

The resource availability vector $\boldsymbol{\mathcal{R}}^*$ can be thought of as a modified version of the supply vector $\mathbf{K}$, where the availability of resource α is reduced from K_α by a factor of $\sum_j \mathcal{F}_{j\alpha} n_j^*$. Therefore alignment between $\boldsymbol{\mathcal{G}}_I$ and $\mathbf{K}$ results in a contribution to the invasion fitness of strain I that is independent of the other strains in the community—this contribution to the invasion fitness is a form of *general fitness*, here characterized by $U_i \equiv \sum_\alpha \mathcal{G}_{i\alpha} K_\alpha$, which captures the average performance of strain i against all of the D resources at the rate that they are supplied by the environment, and thus scales with the magnitude of the resource supply vector. Therefore ξ_I for an invader has a community-independent contribution proportional to the general fitness U_I as well as a community-dependent contribution from the ecological interactions, which set the resource abundances at the fixed point. We refer to the community-dependent part of the invasion fitness as the *drive*, which will be defined more precisely for the models we introduce in Section III.

It is useful for analysis to extend the definition of the invasion fitness to strains already in the community: for $i \in \mathcal{C}$ we define ξ_i as the growth rate strain i would have if it were removed from the community and then reintroduced at low abundance. This quantity arises naturally in the mean field description of ecological models (Section IV and Appendix F) and we will sometimes refer to it as the *bias*, consistent with previous work [23]. The bias is therefore an invasion fitness appropriately defined for strains both inside and outside the community. Differences between biases of different strains come from both the general fitnesses and the drives, and such differences are responsible for the differential growth and extinction of strains.

As we show in Appendix C, in a randomly assembled community the *selective differences*—related to differences between general fitness terms U_i—are *first order* in the phenotypic differences, proportional to $\frac{1}{K_{\text{tot}}} \sum_\alpha \mathcal{G}_{i\alpha} K_\alpha - \langle \mathcal{G} \rangle$,

while the variations in the drives are second order, proportional to both $\frac{1}{K_{\text{tot}}} \sum_\alpha \mathcal{G}_{i\alpha} K_\alpha - \langle \mathcal{G} \rangle$ and $\frac{1}{N} \sum_i \mathcal{F}_{i\alpha} n_i^* - \langle \mathcal{F} \rangle$. We have defined $K_{\text{tot}} \equiv \sum_\beta K_\beta$, and weighted averages $\langle \mathcal{G} \rangle = \frac{1}{N K_{\text{tot}}} \sum_{i,\alpha} \mathcal{G}_{i\alpha} n_i^* K_\alpha$ and similarly for $\langle \mathcal{F} \rangle$. Therefore the selective differences comprise the dominant contribution to the width of the bias distribution, larger by a factor of $\sqrt{D}$ than the contribution from the drives. For large D the selective differences determine both the ecological and early-stage evolutionary dynamics and can, at least initially, strongly limit the diversity. Previous work [36] has forced the contributions to be comparable by scaling the mean and variance of $\mathcal{G}$ and $\mathcal{F}$ as $1/D$, while in fact it is their mean and *standard deviation* that should be the same order in a biologically-motivated model. We will later show that in a long-evolved community, the variations in the general fitnesses become much smaller, indeed comparable to the variations in the drives. However—importantly—this feature must be shown rather than assumed.

C. Eco-evolutionary dynamics

To explore the eco-evolutionary dynamics, we choose $D = 50$ resources with $\kappa = 0.4$ chosen to be well away from the special value $\kappa = 1$ for which there is a Lyapunov function: we investigate that special case later. Starting from a single strain ($L = S = 1$), Figure 1A shows the ecological dynamics of the strain abundances in Equation 1 (with resources integrated out), punctuated by well-spaced evolutionary events when an invader comes in at the start of each new epoch. When the system has reached an ecological fixed point, a new strain is introduced, drawn randomly from the ensemble of independent strains *conditioned* on being able to invade. The new strain joins the community and perturbs the abundances of the other strains, sometimes causing extinctions, until a new stable fixed point is reached. Thereafter another strain is introduced.

From now on, we focus on the evolutionary dynamics and ignore the ecological transients at the beginning of each epoch that drive some strains extinct and others to a stable fixed point. The changes caused by invading strains can be visualized over longer times by coarse-graining the time resolution from ecological to evolutionary epochs and tracking the fixed point abundances of strains (calculated with the algorithm of Appendix D). Such strain abundance trajectories are shown in Figure 1B and, after much evolution, in Figure 1C. We show only a few of the strains for ease of visualization.

Primary quantities of interest are the extant strain diversity L, the population mean general fitness $\langle U \rangle \equiv \frac{1}{N} \sum_i n_i U_i$ (with angular brackets denoting weighted averages over the community), and the probability that a randomly chosen strain can invade, denoted p_{inv}. In addition to the community level properties, we are interested in the statistical properties of strains including the distributions and temporal correlations of their abundances n_i, their general fitnesses U_i and their evolutionary lifetimes which we call $\mathcal{T}_i$ (see Section IV C for more details).

The evolution of the diversity for $\kappa = 0.4$ is shown in purple in Figure 2A. Starting from a single strain, L initially grows fast and then slows down, appearing to saturate at
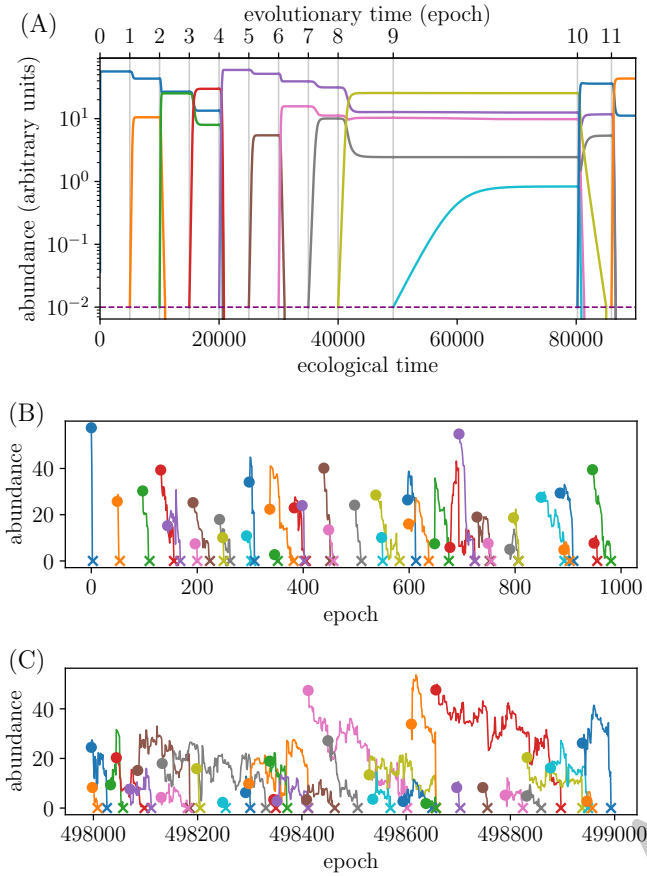
Figure 1. Dynamics of strain abundances on ecological and evolutionary timescales in the consumer resource (CR) model for $\kappa = 0.4$ and $D = 50$ resources. Starting with a single strain, ecological dynamics reach a fixed point between each new strain invasion which starts a new epoch, demarcated by a vertical line. Ecological time on lower axis; evolutionary time, in epochs, on upper axis. (B)–(C) Selected strain abundance trajectories on evolutionary timescales. (B) first 1000 epochs. (C) after a long period of evolution. Turnover of diversity is slower than during initial stages of evolution, but has converged to a steady average rate.

$L \approx 10$ at long times, for which we plot a moving average. As shown in Appendix H for related model, the fluctuations in L are $\mathcal{O}(\sqrt{L})$ resulting from rough independence of strains along with robustness of the community to perturbations (Section IV B). For $\kappa = 0.4$, it thus appears that the evolution converges to a Red Queen steady state. Note that although for these parameters L remains rather small, we later show that it is expected to be of order $\kappa^2 D$, here only 8, in the large D limit of primary interest. The modest $\kappa^2 D$ used here was needed for convergence to steady state in reasonable time.

Further evidence for the Red Queen state for $\kappa = 0.4$ is shown in Figures 2B, where the invasion probability of a new strain, p_{inv}, after initially decreasing rapidly, saturates at long times (albeit to an extremely small value). This behavior can be understood through the general fitnesses. Figures 3A and B show in purple the mean and width, respectively, of the extant U_i, distribution. Initially the mean increases, but then saturates at a value of order
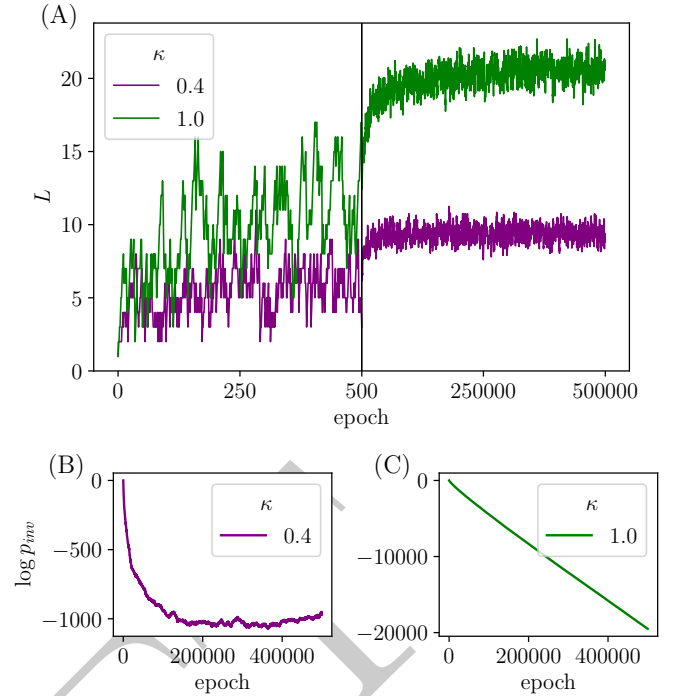

Figure 2. The CR model reaches a steady state for symmetry parameter $\kappa = 0.4$, but not for perfect symmetry, $\kappa = 1$. (A) Community diversity L as function of evolutionary time, starting from a single strain with $D = 50$. First 500 epochs on stretched scale; thereafter as moving average over windows of 1000 epochs. L increases, then fluctuates around steady state for $\kappa = 0.4$, but creeps up for $\kappa = 1$. (B)–(C) The invasion probability decreases initially, then reaches steady state for $\kappa = 0.4$, but continues to decrease for $\kappa = 1$. The absurdly small p_{inv} is an artifact of the gaussian distribution of phenotypes (Section IV E).
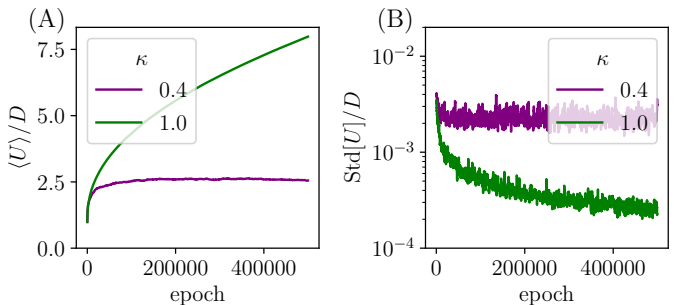
Figure 3. (A)–(B) Mean and standard deviation (weighted by abundances) of extant distribution of generalized fitness, U_i, over evolutionary time for $\kappa = 0.4$ and 1 with $D = 50$. Distribution stabilizes for $\kappa = 0.4$. For $\kappa = 1$, mean continues increasing with std decreasing.

D with the strains all having similar U_i. The evolution therefore pushes far out into the tail of the general fitness distribution, making it very hard for other strains to invade. We shall show in Section IV E that the width of the distribution of extant general fitnesses, scaled by the mean consumption rate of the resources—which is what enters in the growth rate of the strains (Appendix C)—shrinks during evolution until it is similar to the width of the dis-

tribution of drives. At this point the community-derived and intrinsic parts of the bias contribute similarly to bias differences between strains. Once this occurs, the distribution of extant U_i stops pushing higher and we reach a Red Queen evolutionary steady state with a roughly steady rate of turnover.

In an evolutionary steady state (when it occurs) we denote *averages over time* by overbars. Because of the turnover of strains, in steady state temporal averages should be equal to averages over the ensemble of strains.

Turnover of strains: In steady state, one new strain invades in each epoch and on average, one extant strain goes extinct. This means that a strain remains in the community for an average of $\overline{L}$ epochs and the rate of turnover of strains is $1/\overline{L}$ per epoch. This is seen in Figure 1 where at early times when L is small, strains do not survive long, while in the long-time steady state with larger $L \approx \overline{L}$ they survive longer on average even though there are still strains that go quickly extinct, as we discuss in Section IV C.

The continual turnover of strains together with the plateauing of p_{inv} mean that in the Red Queen phase, strains that successfully invade and later go extinct will have a substantial chance of being able to reinvade once the community has sufficiently turned over and lost memory of its state at the extinction time of the strain. Each strain keeps "memory" of its general fitness which, because it was able to invade earlier, must be similar to that of the other strains in the steady state.

While these simulations provide strong evidence for a Red Queen steady state for $\kappa = 0.4$, quantitatively they are problematic. It is apparent from Figure 2B that for the ensemble of strains we have used, the invasion probability of a new invader is unrealistically small in the steady state ($p_{inv} \approx 10^{-400}$). In Section III A we will introduce a family of models in which p_{inv} at steady state can be much larger, and in Section IV E we analyze how the choice of the $\mathcal{G}$ and $\mathcal{F}$ ensembles affects the invasion probability in steady state, and show that the very small $\overline{p_{inv}}$ observed above is an artifact of the gaussian phenotype ensemble we have used for numerical convenience.

Totally symmetric ensemble: We now show that the perfectly-symmetric case, $\kappa = 1$, which possesses a Lyapunov function, Λ, behaves very differently from the $\kappa = 0.4$ ensemble for which there is no Lyapunov function. From the green trajectory of the diversity in Figure 2A, it is difficult, based only on the dynamics of L, to know whether a steady state obtains by the end of the simulation. However, invading strains must increase Λ and there is no absolute upper bound of Λ in the ensemble of all possible invaders. Therefore even if L saturates, there cannot be a steady state. The larger Λ gets, the more unlikely that adding new strains can increase it further—thus the probability of invasion, p_{inv} must steadily decrease, as indeed seen in Figure 2C. Concomitantly, the mean general fitness pushes further and further into the tail of its distribution as shown by the green trajectory in Figure 3A, while the standard deviation of the U_i decreases (green curve Figure 3B) due to the steepening of the distribution around $\langle U \rangle$ (see Section IV E).

Because of the Lyapunov function, we conclude that there is no Red Queen steady state for $\kappa = 1$. It is in-

CR	consumer resource model (Equation 1)
LV	Lotka-Volterra model (Equation 5)
ν_i	fractional abundance of strain i; $\sum_i \nu_i = 1$
S	number of strains in the initial pool
L	number of extant strains in the community
D	number of interaction-mediating resources
Q	niche parameter of LV model; $\mathrm{E}[V_{ii}] = -Q$
$\mathbf{K}$	vector of resource supply rates
$\mathcal{G}$	growth rate matrix
$\mathcal{F}$	feeding rate matrix
G	matrix of growth rate differences; mean 0
F	matrix of consumption rate differences; mean 0
V	pairwise strain interaction matrix
κ	correlation between growth and feeding; $\kappa \in [0,1]$
γ	symmetry parameter for LV model; $\gamma \in [-1,1]$
U_i	general fitness; $U_i = \sum_\alpha \mathcal{G}_{i\alpha} K_\alpha$
p_{inv}	probability of a new variant successfully invading
ξ_i	invasion fitness = bias
ζ_i	drive = community-dependent component of bias
s_i	general fitness = intrinsic component of bias
Υ	community mean "fitness"; $\Upsilon = \sum_i \nu_i(s_i + \sum_j V_{ij}\nu_j)$
Σ	scale of s distribution
$\Delta(\Upsilon)$	scale of s distribution around $s \approx \Upsilon$ (Equation 18)
ρ	correlation between parent and mutant
$\overline{A}$	evolutionary time average of A
$\langle A \rangle$	community average of A (unless otherwise stated)
$\hat{A}$	value of A at the top-down assembled transition

Table I. Definitions of variables and terms.

teresting to note that, in spite of the gradual increase of Λ, L stays far away from its upper bound of $D = 50$. Indeed, L stays below its maximum value of $D/2$ for a randomly *assembled* community with $\kappa = 1$ [36]. However it is not clear whether under evolution L might eventually reach or exceed this bound.

Evolution via mutations: So far we have considered independent invaders in the CR model. In Appendix B we study the evolutionary dynamics when invaders are correlated with a parent strain chosen with probability proportional to its abundance in the community. Indexing the parent by P and the mutant by M, we introduce a parameter $\rho \in [0,1)$: the elementwise Pearson correlation coefficient between $\mathcal{G}_P$ and $\mathcal{G}_M$ and between $\mathcal{F}_P$ and $\mathcal{F}_M$. Given a parent, we draw mutants from this ensemble (while also preserving correlation κ between $\mathcal{G}_M$ and $\mathcal{F}_M$), therefore the correlation between U_P and U_M is also ρ. Our main finding is that for ρ sufficiently far from 1—i.e. mutations of substantial effect—the eco-evolutionary phenomenology is similar to the independent invader limit, $\rho = 0$: This bolsters our confidence in basing qualitative conclusions on results from independent invaders rather than true evolution.

III. SIMPLIFIED MODELS

The saturation of the community-mean general fitness in the Red Queen eco-evolutionary steady state for $\kappa < 1$ is a manifestation of continual evolution in an ecological system. We seek simpler—but still robust—models that display such a Red Queen steady state, but which allow us to separate out and independently tune the roles of ecological interactions and general fitnesses. The two related models that we shall study both have a generalized Lotka-Volterra form, but with different ensembles for the interaction matrix between strains. The first model is a linearized version of the CR model with the resource dynamics integrated out, yielding a maximum rank D for the strain interaction matrix. The second model is of a more standard Lotka-Volterra type, with direct interactions drawn from some ensemble and diagonal "niche" interactions enabling substantial diversity to coexist. For some ranges of parameters, both of these models exhibit a Red Queen phase characterized by constant turnover of biodiversity. In addition they allow us to isolate the effects of the general fitnesses on the eco-evolutionary dynamics.

A. Linearized resource model

In the Red Queen steady state of the CR model, the mean of $\sum_j \mathcal{F}_{j\alpha} n_j^*$, the total consumption rate of resource α, is large compared to its variations (Figure 14A), suggesting an expansion in these variations. In addition, the total population of the ecosystem $N = \sum_j n_j^*$ is roughly constant on evolutionary timescales (Figure 14B), due to the large mean of $\mathcal{G}$ and $\mathcal{F}$ relative to their variations—this motivates us to separate out the mean part of these matrices from their variations, and introduce zero-mean residual interaction matrices G and F. We additionally work in terms of fractional abundances $\nu_i = n_i/N$ rather than absolute abundances, and integrate out the resource dynamics since these do not affect the fixed point.

Expanding in the *differences* between resources and between strains yields general fitnesses at first order and interactions at second order (Appendix C). The simplified model that we introduce, and refer to hereafter as the *linearized resource model*, or simply the linearized model, is

$$\frac{d\nu_i}{dt} = \nu_i \left(s_i + \sum_{\alpha=1}^{D} G_{i\alpha} R_\alpha - \Upsilon(t) \right) \quad (4a)$$

$$R_\alpha = -\sum_{j=1}^{S} F_{j\alpha} \nu_j. \quad (4b)$$

The term $\Upsilon(t)$ is a Lagrange multiplier, equal to $\sum_i \nu_i(s_i + \sum_\alpha G_{i\alpha} R_\alpha)$, which incorporates the average effects of growth, death, and resource depletion, and keeps the population constant with $\sum_i \nu_i = 1$. Here, the per capita growth rate of each strain is a linear function of the abundances of all the other strains, with a quadratic part from Υ that is the same across all the strains and includes the average general fitness. Differences in general fitness are parametrized by s_i, the *selective differences* between strains, which play

the role of the U_i in the CR model. As the means have been subtracted off, the residual part of the resource vector $\mathbf{R}$ can have components which are negative. $\mathbf{R}$ then measures *deviations* of resource availability away from some positive baseline.

The structure of Equation 4 is that of a generalized Lotka-Volterra model or system of replicator equations (discussed further in Section III B), where the interaction matrix between S strains, denoted by V_{ij}, has a structure inherited from the G and F matrices via $V = -GF^{\intercal}$, with G and F are $S \times D$ matrices. We take each element of G and F to be gaussian distributed with mean zero and unit variance, with no correlations within G or F. Elementwise correlations between G and F are again parameterized by $\kappa \in [0, 1]$ via $\mathrm{E}[G_{i\alpha} F_{j\beta}] = \kappa \delta_{ij} \delta_{\alpha\beta}$.

The off-diagonal entries of V have mean zero and standard deviation $\sqrt{D}$ while the diagonal entries have mean $-\kappa D$ and standard deviation $\sqrt{(1+\kappa^2)D}$. The parametrically-larger (for large D) diagonal entries thus endow the interactions with sufficient *niche* structure to stabilize the dynamics. (Note that $\kappa < 0$ would result in unstable behavior with the diagonal entries of V having positive mean.) As in the CR model, the number of resources D sets the dimensionality of phenotype space, with the phenotype of strain i comprised of s_i along with the ith rows of G and F. The maximum number of coexisting strains at a stable fixed point of the linearized model is equal to D and will generally be less than—though proportional to—D.

When $\kappa = 1$, Υ is a convex Lyapunov function for Equation 4. If additionally $s_i = 0$, we have $\Upsilon = -\sum_\alpha (\sum_i G_{i\alpha} \nu_i)^2$ which is nonpositive. For nonzero s_i, the Lyapunov function gets an extra piece $\langle s \rangle \equiv \sum_i \nu_i s_i$ and can be of either sign, though it remains a convex function of the $\{\nu_i\}$.

We will take the general fitnesses s_i to be drawn from a distribution $p(s)$, which for most of this paper is gaussian with mean 0 and variance Σ^2. In Section IV E we show that the value of Σ determines the early evolutionary behavior and the value of p_{inv} at steady state, with the shape of the large-s tail of $p(s)$ playing a controlling role. As the community evolves and the distribution of extant s_i pushes into the tail of $p(s)$, the effective Σ—the variance of the extant s_i—decreases. Therefore tuning Σ to be small *a priori* mimics the effects of *prior evolution* and allows us to study communities that have been evolving for a long time, without the long initial evolutionary transient.

As discussed previously in Section II B, for the CR model in Equation 1, the width of the U_i distribution is a factor of $\sqrt{D}$ larger than the width of the drive distribution. In the linearized model, this corresponds to $\Sigma = \mathcal{O}(\sqrt{D})$, since the distribution of the drives in a diverse community has $\mathcal{O}(\sqrt{D/L}) = \mathcal{O}(1)$ width. However, in the Red Queen steady state, as we shall see, the first and second order—in $1/\sqrt{D}$—components of the growth rate of each strain become comparable to one another. Thus we mostly consider $\Sigma = \mathcal{O}(1)$, and sometimes set $\Sigma = 0$, which shortens the transient period of evolution before the steady state is reached—and enables analysis—but does not change our conclusions. In Section IV E we also discuss non-gaussian distributions of the s_i, which affect quantitative properties

but not the basic phenomenology.

B. Lotka-Volterra model

In addition to the linearized model of Equation 4, we have studied the generalized Lotka-Volterra (LV) model (or replicator equations) [48] which take the form

$$\frac{d\nu_i}{dt} = \nu_i \left(s_i + \sum_{j=1}^{S} V_{ij}\nu_j - \Upsilon(t) \right),\qquad(5)$$

with $\Upsilon(t) = \sum_i \nu_i(s_i + \sum_j V_{ij}\nu_j)$. The LV model has similar structure to the linearized model of Section III A; the difference lies in the statistics and parameterization of the interaction matrix V. Instead of the $V = -GF^\intercal$ with random G and F of the linearized resource model, in this LV model the V_{ij} are themselves gaussian random variables with $\mathrm{E}[V_{ij}] = -Q\delta_{ij}$, $\mathrm{E}[V_{ij}^2] = 1 + \gamma\delta_{ij}$, and $\mathrm{E}[V_{ij}V_{ji}] = \gamma$ for $i \neq j$ and other conveniences zero. Here $\gamma \in [-1, 1]$ is the symmetry parameter tuning between predator-prey-like interactions for $\gamma < 0$ and competitive interactions for $\gamma > 0$: a rough correspondence with the linearized model for positive γ is $\gamma \sim \kappa^2$. The negative diagonal term $-Q$ is a "niche" parameter that measures the strength of self-interaction relative to inter-strain interaction, playing a similar role to $-\kappa\sqrt{D}$ in the linearized model. Our parameterization of the LV equations differs from a common parameterization in which interactions have nonzero mean [19], here the V matrix captures *variations* in interaction magnitudes, with Υ taking the role of the mean interaction strength [23]. As in the linearized model, when the interactions are symmetric ($\gamma = 1$), Υ is a Lyapunov function of the dynamics, though in contrast to the linearized model [35], it can become a nonconvex function of the $\{\nu_i\}$ for large enough S [37].

The LV model has a similar ecological phase diagram to the CR model and linearized models, with a globally attracting fixed point of the dynamics when S is sufficiently small. The maximum L at a stable uninvadable fixed point scales as Q^2 with a γ-dependent coefficient [20]. Note that the scale of the drive distribution is $\mathcal{O}(1/\sqrt{L}) = \mathcal{O}(1/Q)$ in the LV model, as opposed to $\mathcal{O}(\sqrt{D/L}) = \mathcal{O}(1)$ in the linearized model, since the overall scale of the entries in the LV model is smaller by a factor of $\sqrt{D}$ than in the linearized model.

In the LV model, unlike in the linearized and CR models, strains cannot be characterized by finite-dimensional phenotypes. Although one can attempt to define a phenotype of strain i by s_i along with its corresponding row and column of V, this phenotype depends on the rest of the community and thus changes—or grows in dimension—as the community turns over.

C. Properties of invaders

Having defined the simplified ecological models of interest, we will now specify the process by which they evolve. We first discuss general facts about invaders introduced into the community, and then define the specific parameterization of invaders that we have used for simulations.

As in the CR model, whether a new strain I successfully invades the extant community or not is determined by its invasion fitness or bias, now written as $\xi_I = s_I + \zeta_I - \Upsilon$ with $\zeta_I = \sum_j V_{Ij}\nu_j$ in the LV model and $\zeta_I = -\sum_\alpha G_{I\alpha}R_\alpha$ in the linearized model. Splitting the bias up into the general fitness piece s_I, the Lagrange multiplier Υ and the drive ζ_I is useful since it separates the intrinsic and extrinsic contributions to the bias. Strains with positive bias can invade the community and reach abundance proportional to their bias, whereas strains with negative bias go extinct. As in the CR model, we sample properties of the invader I conditional upon $\xi_I > 0$ [47]. A strain's extinction occurs when its abundance goes to 0 in the ecological dynamics or, equivalently, when its bias goes negative.

In the LV model, the lack of phenotypes makes attempts to parameterize correlations between parent and mutants *ad hoc* for general γ [43], so we confine our study to the case of independent invaders. An invader strain I is generated by drawing a new row and column of V from the original gaussian ensemble with $\mathrm{Corr}[V_{Ii}, V_{jI}] = \gamma\delta_{ij}$, and by drawing general fitness s_I from a gaussian distribution with mean 0 and variance Σ^2. Therefore the invasion probability of an independent invader is $p_{inv} = 1 - \Phi[\Upsilon/\sqrt{\sum_i \nu_i^2 + \Sigma^2}]$ where Φ is the standard normal CDF.

In models with phenotypes, such as the CR or linearized model, there is a natural way to generate mutants through small changes to the phenotype of an extant parent strain. This procedure creates a mutant whose interactions with all other strains are similar to those of its parent. In our gaussian ensemble, the probability of any parent generating a mutant that can invade is the same—therefore we can, in an unbiased way, simply select a parent with probability proportional to its abundance.

The parameter $\rho \in [0, 1)$ describes the correlation between parent P and mutant M. The mutant phenotype is drawn from an ensemble that preserves $\mathrm{E}[G_{M\alpha}G_{P\beta}] = \mathrm{E}[F_{M\alpha}F_{P\beta}] = \rho\delta_{\alpha\beta}$, $\mathrm{E}[G_{M\alpha}F_{M\beta}] = \kappa\delta_{\alpha\beta}$, and $\mathrm{Corr}[s_M, s_P] = \rho$. We fix the desired correlations by drawing the phenotype of M as

$$G_{M\alpha} = \rho G_{P\alpha} + \sqrt{1 - \rho^2} Z_\alpha \qquad (6a)$$

$$F_{M\alpha} = \rho F_{P\alpha} + \sqrt{1 - \rho^2}(\kappa Z_\alpha + \sqrt{1 - \kappa^2}Z'_\alpha), \qquad (6b)$$

$$s_M = \rho s_P + \Sigma\sqrt{1 - \rho^2}Z'' \qquad (6c)$$

where the $\{Z_\alpha\}$, $\{Z'_\alpha\}$ and Z'' are uncorrelated standard normal random variables. This choice ensures that the elements of the interaction matrix $V = -GF^\intercal$ have correlations $\mathrm{Corr}[V_{Mi}, V_{Pj}] = \mathrm{Corr}[V_{iM}V_{jP}] = \rho\delta_{ij}$ for $i, j \notin \{M, P\}$. The invasion probability of a mutant drawn from this ensemble is given by

$$p_{inv} = 1 - \Phi\left[\sqrt{\frac{1 - \rho}{1 + \rho}}\frac{\Upsilon}{\sqrt{|\mathbf{R}|^2 + \Sigma^2}}\right] \qquad (7)$$

where Φ is the standard normal CDF. In this paper we restrict study to large effect mutations, with $\sqrt{1 - \rho^2} = \mathcal{O}(1)$. The behavior as $\rho \to 1$, corresponding to small effect mutations, is both interesting and biologically relevant: we will discuss it in depth in future work.
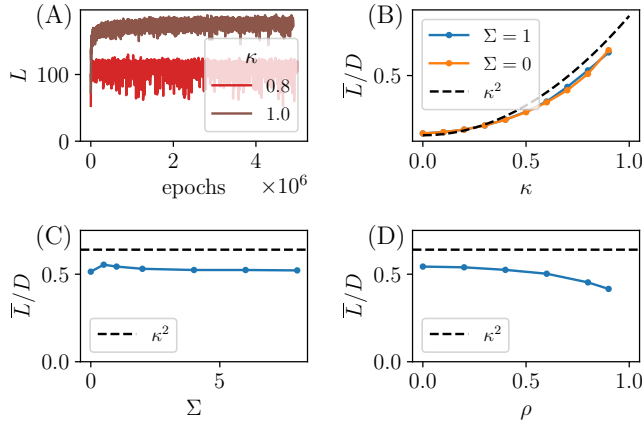
Figure 4. Dynamics and steady-state of diversity in linearized model as function of κ, Σ and ρ. Default values $D = 200$, $\kappa = 0.8$, parent-mutant correlation $\rho = 0$, width of distribution of selective differences $\Sigma = 1$, but each varied as specified per panel. (A) Dynamics of L over evolutionary time for $\kappa = 0.8, 1$. Although L appears to be reaching steady state for both parameters, for $\kappa = 1$, Υ and p_{inv} indicate it is not (Figures 5 and 6). (B) Steady state diversity. $\overline{L}$ normalized by its maximum, D, as function of κ, for $\Sigma = 0$ and 1. Dotted line shows maximum diversity in *top-down assembled* community; in Red Queen phase, diversity is lower. (C) $\overline{L}$ depends only weakly on Σ. (D) L depends weakly on ρ for ρ not too near unity.
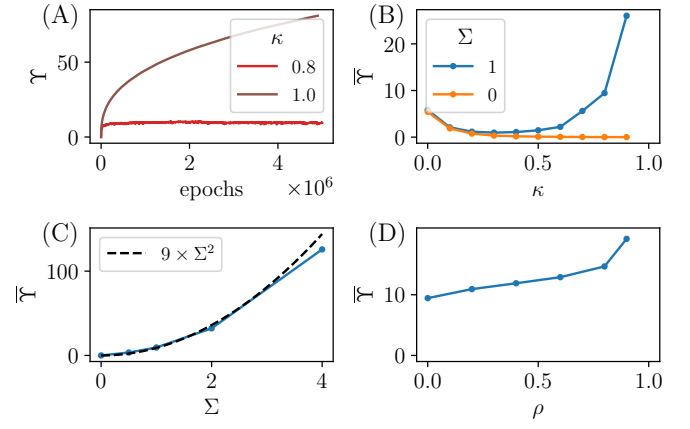

Figure 5. Dynamics and steady state of Lagrange multiplier Υ —loosely community "fitness"—as a function of parameters for linearized resource model with default values as in Figure 4. (A) Dynamics: for $\kappa = 1$, Υ (Lyapunov function in this case) is ever-increasing. For $\kappa = 0.8$, Υ reaches steady state $\overline{\Upsilon}$. (B) $\overline{\Upsilon}$, as function of κ for $\Sigma = 0$ and 1. Only for $\Sigma > 0$ does the phenotype push into the tail of its distribution indicated by large Υ. Simulations are less reliable for small κ because L is smaller (see Appendix D). For $\kappa = 0.9$, $\Sigma = 1$, steady state not yet reached by end of simulation. (C) Dependence on Σ. Dotted line is rough fit to predicted form, $\overline{\Upsilon} \sim \Sigma^2$. At $\Sigma = 4$, Υ has not yet saturated. (D) Weak dependence of $\overline{\Upsilon}$ on ρ.

D. Evolution in the linearized resource model

The motivation for introducing the linearized model is to capture the phenomenology of the CR model— a Red Queen state with continual turnover but roughly-constant general fitness and p_{inv}—in a simple setting where the effects of the general fitnesses can be separated out from the interactions, and with a shorter transient period before the steady state is reached. We begin by describing the basic phenomenology of evolution in the linearized model of Equation 4, with the evolutionary dynamics of Equation 6. In general, the evolutionary dynamics are parameterized by D, κ, ρ, Σ and S, the size of the initial pool of strains. However S only plays a role in the transient behavior: the evolution converges to the same state independent of the initial diversity as illustrated in Section III E for the LV model.

In Figure 4A we plot the dependence of L on evolutionary time for two values of κ: 0.8 and 1. (We have here chosen $\kappa = 0.8$ to yield higher diversity than the $\kappa = 0.4$ used for the CR model.) Starting from a single strain, there is a transient during which the diversity builds up, after which both trajectories flatten out. Figure 5A shows the trajectories of the Lagrange multiplier, Υ—analogous to the mean general fitness $\langle U \rangle$ in the CR model, which clearly showed the difference between the behavior with and without a Lyapunov function. For $\kappa = 0.8$, saturation of Υ indicates a Red Queen steady state as this implies saturation of p_{inv}. By contrast, for $\kappa = 1$, Υ continues to increase and p_{inv} concomitantly decreases: there is no Red Queen phase.

To investigate the Red Queen phase in the absence of a Lyapunov function, we study the long-time averages $\overline{L}$, $\overline{\Upsilon}$

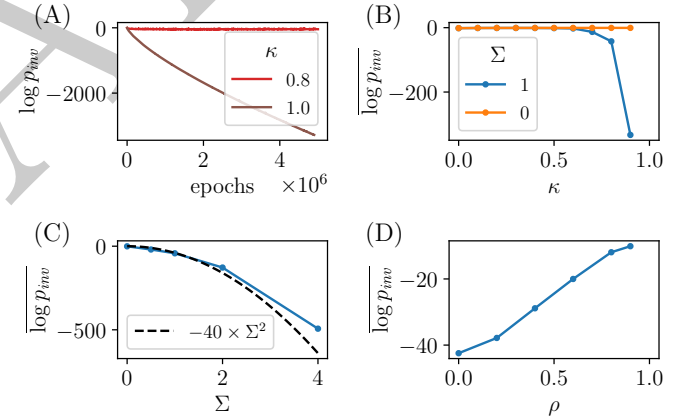Figure 6. Dynamics of invasion probability, and mean logarithm of invasion probability in steady state, $\overline{\log p_{inv}}$, as function of parameters with default values as in Figure 4. (A) Dynamics: p_{inv} saturates for $\kappa = 0.8$ but continues to decrease for $\kappa = 1$. (B) Dependence on κ. (C) Dashed line is rough fit to predicted $-\overline{\log p_{inv}} \sim \Sigma^2$, by analogy with assembled community. (Equation 8b). Simulations for $\Sigma = 4$ have not reached steady state and $\overline{\log p_{inv}}$ is overestimated (see Appendix J). (D) Invasion probability increases with ρ for ρ far from 1.

and $\overline{\log p_{inv}}$, for ranges of the parameters κ, Σ, and mutant-parent correlation ρ. (Note that the *typical* p_{inv} of interest is better characterized by $\overline{\log p_{inv}}$ than $\log \overline{p_{inv}}$ as p_{inv} can fluctuate by large amounts especially when Σ is large as discussed in Section IV E.) Figures 4, 5 and 6 display these quantities across a range of κ, Σ, and ρ. The important conclusion from these simulations is that the Red Queen phase obtains across a range of parameters for $\kappa < 1$, and
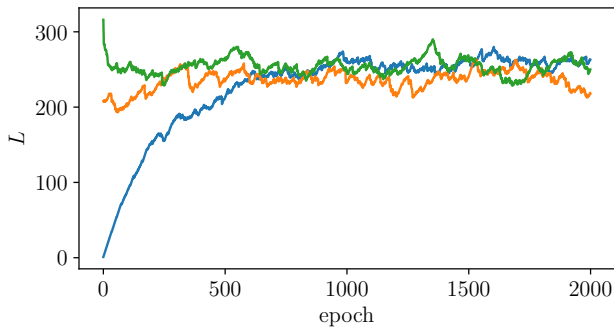
Figure 7. Same steady state is reached in LV model for a range of initial strain pool sizes S which determine initial diversity. All initial conditions tend toward same Red Queen phase after short initial transient. Parameters of LV model $Q = 20$, $\gamma = 0$, $\rho = 0$, $\Sigma = 0$.

that parameter choices, though affecting the length of the pre-steady state transient and the qualitative properties of the steady state, do not eliminate the Red Queen phase. Therefore the Red Queen state can be qualitatively understood by studying just a few values of the relevant parameters. The value of D sets the overall scale of L but, as long as it is large, does not matter much—we therefore choose a large $D = 200$ and $\kappa = 0.8$ for our simulations. This choice yields a large number (roughly $D/2$) strains surviving in the Red Queen phase yet keeps away from the non-generic $\kappa = 1$ behavior.

In Figure 4A, the observed $\overline{L}$ is compared with the curve $\kappa^2 D$ which is the maximum L in a top-down assembled community with F and G drawn from the gaussian ensemble for the linearized model (Section IV A). We see that $\overline{L}$ from simulations is less than this naive upper bound as discussed in Section IV B. In Section IV E we give scaling arguments for how $\overline{\Upsilon}$ and $\overline{\log p_{inv}}$ at steady state depend on κ and Σ: some of these scalings are compared with simulations in Figures 5 and 6.

E.   Evolution in the LV model

Given the apparent similarities between the CR and linearized models, it is natural to ask how the eco-evolutionary phenomenology in the random Lotka-Volterra model resembles or differs from these: in particular whether the difference in the ensemble of the interaction matrix between the LV and linearized models matters. In Appendix E we investigate evolution—by independent invaders—in the LV model across a range of γ and Q, the symmetry and niche parameters defined in Section III B. For a range of γ we find a Red Queen phase analogous to those of the CR and linearized model, with the product of Q^2 and a function of γ determining the number of strains in the steady state; we have also checked that a Red Queen phase obtains for nonzero Σ. Similarly to in the linearized model, L in the Red Queen phase fluctuates around a mean value $\overline{L}$ that is less than the maximum L from a top-down assembled community. The gap between the realized $\overline{L}$ and the assembled upper bound as a function of γ is shown in Figure 15A. This gap diverges as $\gamma \to -1$, meaning that in this regime,

the Red Queen steady state is much less diverse than the largest possible community assembled top-down. In contrast, in the linearized model, the gap is bounded over the whole range of κ. The distinction is not surprising: there is no regime of the linearized model that corresponds to the negative cross-diagonal correlations in V that characterize the LV model with negative γ.

The fact that a top-down assembled community can be substantially larger than the evolved community in the LV model allows us to check that the Red Queen phase obtains across a range of initial number of strains, S. In particular, since the top-down assembled community can have $L > \overline{L}$, there are situations in which there is a naturally-findable stable ecological community that is unstable to evolution, with L decreasing as the evolving community converges to the Red Queen phase. Figure 7 shows this behavior in the LV model for $\gamma = 0$ and $\Sigma = 0$, with L approaching $\overline{L}$ from both above or below, depending on the S of the initial community. Therefore, even in simple models, the question of whether evolution increases or decreases diversity from a top-down assembled initial condition has no general answer [28, 30].

Breakdown of Red Queen phase: Although we find a robust Red Queen phase for a range of γ and Σ in the LV model, closer investigation of the effect of γ reveals an interesting feature. The evolutionary dynamics in the LV model change sharply with γ, appearing to exhibit a discontinuous transition around $\gamma_c \approx 0.65$ for $\Sigma = 0$. For $\gamma < \gamma_c$ a Red Queen phase obtains, while for $\gamma > \gamma_c$ we find a phase in which the rate of successful invasion is ever-slowing and a few strains rise to high abundance. We therefore name this phase the "oligarch" phase of eco-evolutionary dynamics— a counterpart to the Red Queen's monarchy. The oligarch phase is discussed further in Section V, and we leave its fuller analysis for future work. For now, we note that there does not appear to be an oligarch phase in the linearized or CR models: the Red Queen steady state obtains for all $\kappa < 1$, albeit with $\overline{\Upsilon}$ and $\overline{\langle U \rangle}$ respectively diverging as κ approaches 1 and evolutionary transients correspondingly lengthening.

IV.   ANALYSIS

In order to understand the phenomenology of the Red Queen phase described above, it is useful to first discuss the *top-down assembled* community of the linearized model in which all strains are brought together at once with dynamics determined by their ecological interactions. An understanding of the ecological phase diagram in this setting is informative for analyzing the Red Queen phase, and gives us an idea of how $\overline{\Upsilon}$ and $\overline{\log p_{inv}}$ scale with various parameters in the Red Queen phase.

To analyze the eco-evolutionary dynamics, we will introduce a particular limit of the linearized model which reduces to the LV model with $\gamma = 0$. With the further simplifications of independent invaders and $\Sigma = 0$—yielding the simplest ensemble of the LV model—we can exactly solve for the steady state of the evolutionary dynamics in the limit of large diversity. Key quantities are the distribution of strain abundances in an evolved community—strikingly
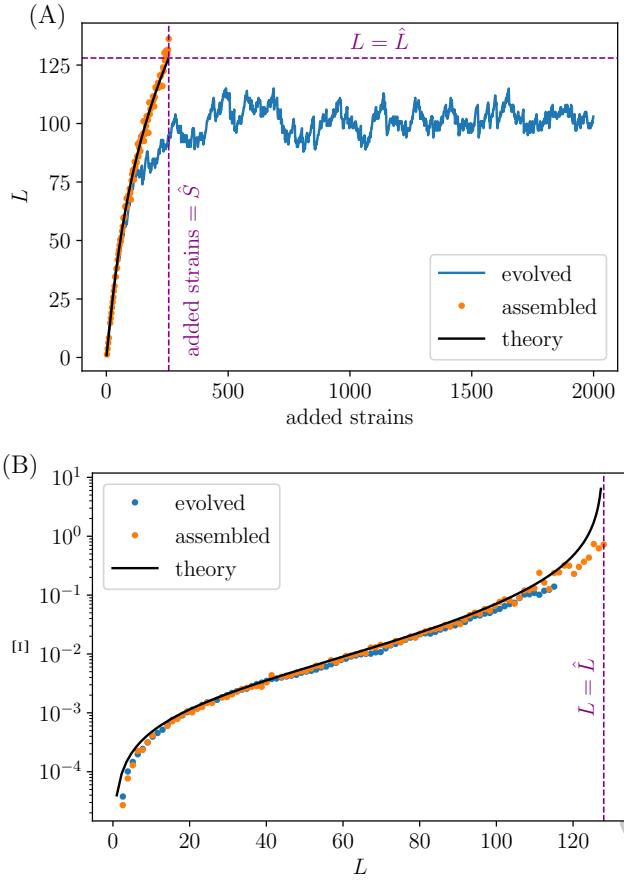
Figure 8. Maximum diversity and stability to evolutionary perturbations comparing top-down assembled (yellow, black) with evolved (blue) communities, with data for small L taken during build-up from a single strain. Linearized model with $\kappa = 0.8$, $D = 200$, and $\Sigma = 0$. (A) Number of surviving strains as function of number of strains added. For assembled community, strains added all at once; for evolved community, one at a time. Maximum diversity in evolved case is less than saturated diversity, $\hat{L}$ (horizontal dashed line) of assembled communities, reached with $\hat{S} = 2\kappa^2 D$, (vertical dashed line). (B) Fragility Ξ, a community-wide measure of susceptibility to perturbations, versus L. Fragility diverges at transition between stable and unstable phases (vertical dashed line), which is not reached in evolved communities.

different than a top-down assembled community—and temporal correlation functions of strain abundances in evolutionary time.

Then we add in the effects of selective differences, presenting general scaling arguments for the dependence of $\overline{L}$, $\overline{\Upsilon}$ and $\overline{\log p_{inv}}$ on Σ, which build on the intuition gleaned from the assembled communities. We generalize to different distributions, $p(s)$, and show how the primary pathologies of the gaussian distribution can be alleviated.

A. Assembled communities

In a top-down assembled community of the linearized model, all the strains are initialized at positive abundance, and are governed by the deterministic dynamics of Equation 4 without a sharp extinction cutoff. The cavity method [49] can be used to analyze both the statics and dynamics of the assembled community with S strains in total and D resources, in the limit of large S and D with the ratio S/D constant. The spirit of the cavity method—in the static case—is to consider a large community to which one adds, separately, one new strain and one new resource: these each act as small perturbations to the fixed point of the community. The abundance of the new strain is determined by its invasion fitness, modified by the community-wide response to its introduction, and the new resource is treated similarly. One then enforces self-consistency conditions that the statistics of the strains and resources in the community are the same as the statistics of the newly added strain and resource. These conditions rely on the community being assembled with independent interactions, since then the independently-drawn invading strain and resource are from the same ensemble as the rest of the community. By solving the self-consistency equations one can extract abundance distributions of ν_i and R_α, and solve for L/D and Υ as a function of S/D, κ and Σ. Similar results can be obtained for the non-linear CR model [36] and—more simply—the LV model [20].

The cavity method also allows one to derive self consistent equations for the ecological *dynamics* [49], in addition to their fixed point—we use a variant of this method to understand the *evolutionary* dynamics in Section IV C.

Details of the static cavity analysis for top-down assembled communities in both the linearized and LV models are in Appendix F. We find that generically the linearized model exhibits two phases as a function of the ratio S/D. For $\Sigma > 0$, when S/D is small enough, there is a unique stable globally attracting fixed point, uninvadable by the strains that are going extinct: this fixed point is described by the cavity solution and the community diversity is $L/D < \kappa^2$. The abundances of the strains in the stable community are distributed as a truncated gaussian so that there is a constant density of abundances near zero. (Figure 9 shows this feature—and its agreement with the result of the cavity calculation—for the LV model which is similar to the linearized model in this respect.) The strains destined for extinction have negative invasion fitnesses with a distribution that also has a constant density near zero. Thus the assembled community has many close-to-marginal strains.

When the number of the strains assembled, S, becomes larger than a (Σ-dependent) critical value $\hat{S}$, at which $L = \hat{L} = \kappa^2 D$, the cavity solution becomes unstable and chaotic dynamics or multi-stability ensues [21]. We use a hat to denote the values of quantities at this phase transition, and refer to the community at this transition as *saturated*. The number of initial strains needed to achieve this saturated state, $\hat{S}$, diverges as $\kappa \to 1$ as long as $\Sigma > 0$. Therefore for $\kappa = 1$ and $\Sigma > 0$ the cavity solution is always stable. When $\Sigma = 0$ the picture for $\kappa < 1$ remains the same, while at $\kappa = 1$ there is a transition (for $S > 2D$) to a nongeneric "shielded" phase [35] in which $L/D = 1$—the maximum allowed by competitive exclusion—and this state is highly marginal.

With an extinction threshold, even for $S > \hat{S}$ the ecological dynamics will generally reach a stable fixed point with

$L \lesssim \kappa^2 D$. However the resulting community will generally be unstable to re-invasion by some of the strains that had gone extinct. Furthermore, in this super-saturated regime there are multiple possible stable communities, almost all of them invadable by some of the extinct strains. Which community occurs will thus depend on the initial conditions. The question of whether or not stable uninvadable communities are likely to exist for large D is unresolved except for the special $\kappa = 1$ case where there is an absolute maximum of the convex Lyapunov function. However it is believed that even if such fully-stable fixed points exist for $\kappa < 1$, they may attract only an exponentially small fraction of initial conditions [17, 50].

We can glean insights about the properties of an evolved community with $\rho = 0$ from an analysis of the near-saturated assembled community, since after long period of evolution, a large number of strains—many more than $\hat{S}$—will have had a chance to invade the evolved community. Naively, one might think that the end result will be similar to assembling the whole community at once. However as we shall see, the evolved community never becomes saturated in the way that an assembled community can be.

We would like to understand the behavior for $\Sigma \gg 1$ which is what occurs "naturally" (e.g. from the CR model): we thus focus for now on this limit. In the linearized model the saturated assembled community has $\hat{L} = \kappa^2 D$ for all Σ and, for $\Sigma \gg 1$,

$$\hat{\Upsilon} \approx \frac{\Sigma^2 \kappa}{1 - \kappa^2} \tag{8a}$$

$$\hat{S} \sim \exp\left[\frac{\kappa^2 \Sigma^2}{2(1 - \kappa^2)^2}\right]. \tag{8b}$$

The reciprocal of $\hat{S}$ scales with the fraction of surviving strains, and therefore the invasion probability of a single independent invader is of order $D/\hat{S}$.

We conjecture that the *forms* of the scalings of $\hat{\Upsilon}$ and $\log \hat{S}$ with Σ in the assembled community also describe the time-averaged quantities in the Red Queen steady state, i.e. that $\overline{\Upsilon} \sim \hat{\Upsilon}$ and $\overline{\log p_{inv}} \sim -\log \hat{S}$, albeit with unknown coefficients. The scalings in Equation 8, along with the guess that $\overline{L} \sim \hat{L}$, motivate the approximate theory curves in Figures 4, 5 and 6: these are consistent with numerical data from the Red Queen phase. In Section IV E we give an argument that the conjectured scalings of $\overline{\Upsilon}$ and $\overline{p_{inv}}$ with Σ should indeed be expected on evolutionary grounds.

B. Fragility of assembled and evolved communities

In the eco-evolutionary Red Queen phase, ecological communities remain stable, but are different from top-down assembled communities since their interaction statistics are conditioned on evolution and, crucially, they would be *unstable to invasion* by some fraction of previously extant strains that have gone extinct. An indication of this distinction, shown in Figures 4 and 8A, is the diversity of the Red Queen phase: $\overline{L}$ is smaller than in saturated assembled communities, $\overline{L} < \hat{L} = \kappa^2 D$, across the range of κ. When a community is assembled one strain at a time with *permanent* extinctions, it is not possible to reach as high diversity

as reached by top-down assembly. What is responsible for this gap between $\overline{L}$ in the Red Queen phase and $\hat{L}$ from the saturated assembled community? The answer is associated with the relative fragility of the two states.

The transition between phases in the assembled community is marked by the onset of an instability in the cavity solution. This instability entails the divergence of a community-level property that we call the *fragility* and denote by Ξ. The fragility quantifies the variance of the changes in strain abundances in response to a random perturbation to the growth rates of all the strains. Concretely, if for each strain i we add a small independently random term h_i to its growth rate, with $\mathrm{E}[h_i] = 0$ and variance $\mathrm{E}[h_i^2]$ (loosely like applying random fields in magnetic systems), the fragility is defined by $\sum_i \mathrm{E}[\delta \nu_i^2] = \Xi \times \mathrm{E}[h_i^2]$. The fragility can also be written as $\Xi = \sum_i \chi_{ii}^2 \times \mathrm{Tr}[\chi^\intercal \chi]$ where the $\chi_{ij} = \frac{d\nu_i}{d\xi_j}$ are elements of the susceptibility matrix χ of the community. Deep in the stable phase the fragility is small, as small perturbations to strains' growth rates do not cause abundances to reshuffle by much. However as the transition is approached and L nears $\kappa^2 D$, the fragility diverges, indicating a breakdown in the stability of the cavity solution. In Figure 8B we plot the fragility, computed directly from the interaction matrix, for both an evolving and top-down assembled community across a range of L. In both cases, Ξ increases with L, and in the evolving community it plays an important role in the Red Queen phase, as we now discuss.

Each new invader adds a small random part to the biases of the other strains. Invader I, which reaches a stable abundance ν_I, contributes $V_{iI}\nu_I$ to the bias of each strain i. The variance of this random perturbation is $\mathrm{Var}[V_{iI}\nu_I] = \nu_I^2 D$ for the linearized model. This is multiplied by a factor of $\Xi/L \sim 1/L^2$ to yield $\mathrm{Var}[\delta\nu_i]$ of order $1/L^3$ and therefore $\delta\nu_i \sim L^{-3/2}$. For a typical strain with $\nu_i \sim 1/L$, this perturbation will be small. However in a top-down assembled community, the lowest abundance strains will have ν_i smaller than typical by a factor of $1/L$, and $\mathcal{O}(\sqrt{L})$ of these will have abundance less than the magnitude of the perturbation. This implies that the invasion will drive the bias of $\mathcal{O}(\sqrt{L})$ of the lowest abundance strains negative, causing their extinction. Concomitantly, if we do not enforce permanent extinctions, the perturbation will increase the bias of a similar number of barely-extinct strains and these will enter the community. As long as Ξ is not large, this number of extinctions and invasions will be small. As $S \nearrow \hat{S}$, Ξ diverges and the assembled community will be shaken a lot by a single additional strain.

The behavior of an *evolved* community—with permanent extinctions—in response to a small random perturbation is quite different. When the community has just been assembled, if it is close to saturated, a single invader will drive a fraction of the strains with $\nu \leq \mathcal{O}(L^{-3/2})$ extinct. As this is $\mathcal{O}(\sqrt{L})$ strains, L will initially decrease (Figure 7). Under the perturbations from further invasions, the density of low abundance strains will be hollowed out by this process which is like diffusion with an absorbing boundary condition at zero: This results in a density of abundances that goes linearly to zero as $\nu \to 0$ (see Figure 9). The diffusion coefficient of the abundances in evolutionary time is proportional to the fragility of the community, so if Ξ is

large, then the perturbations are amplified and more extinctions result, on average, from each successful invader. As the diversity grows, we expect that (as in the top-down assembled communities), the fragility also grows. When it is large enough that on average one extinction is driven by each invasion, the diversity stabilizes. This must thus occur while the fragility is not too large suggesting—although not implying because of the complexities on conditioning on the evolutionary history—that the Red Queen state will be less diverse than the maximal assembled diversity, $\hat{L} = \kappa^2 D$, as observed in Figure 8A. Note that the ratio $\overline{L}/\kappa^2 D$ depends on the value of κ (Figure 4B); this is presumably due to the dependence of Ξ on κ.

In the special case of perfect symmetry ($\kappa = 1$ for the linearized model and $\gamma = 1$ for the LV model), L approaches $\hat{L}$ at long times despite the lack of a Red Queen steady state. We do not have an explanation for this apparent saturation of diversity.

In recent work, de Pirey and Bunin [51] have observed and studied a similar phenomenon of depletion of low-abundance strains due to continual perturbations to the community from invading strains. These authors looked at a purely *ecological* model but with small migration from a mainland that prevents total extinction of any strain. Species rising up from the brink of extinction play the role of invaders in our evolutionary model. The similarities between this ecological model and our work are discussed in Appendix L. A primary difference is due to the migration-induced boundary condition: As there is a flux of strains both in and out, the steady state distribution of ν does not vanish as $\nu \to 0$ and the boundary condition must be determined self-consistently from the ecological dynamics. In the evolutionary problem we study, more analytical progress is possible due to the simplicity of the extinction boundary condition in contrast to that which arises from migration.

C. Exact solution in simplest case

Having understood some features of the eco-evolutionary phenomenology by comparison with the saturated assembled community, we now proceed to directly analyze the evolutionary dynamics for a particularly simple choice of parameters.

With $\kappa > 0$, there are correlations κ^2 between the cross-diagonal elements of the interaction matrix V, as in the case of the $\gamma > 0$ LV model. These correlations induce complex memory into the evolutionary dynamics since the effect of an invader I on a strain j in the community via V_{jI} feeds back via V_{Ij} onto I. This feedback on I is coherent across the extant strains with magnitude proportional to κ^2 (or γ). Thus the properties of a strain in the community depend on its effects on the other strains over all the epochs since its invasion—in particular on its abundance in these earlier epochs. This fact makes the evolutionary dynamics difficult to analyze (like the ecological dynamics in ref. [23]), as the form of this memory kernel needs to be self-consistently determined along with the dynamics of the strains abundances in evolutionary time. We set this analysis up in Section IV D but do not carry it through.
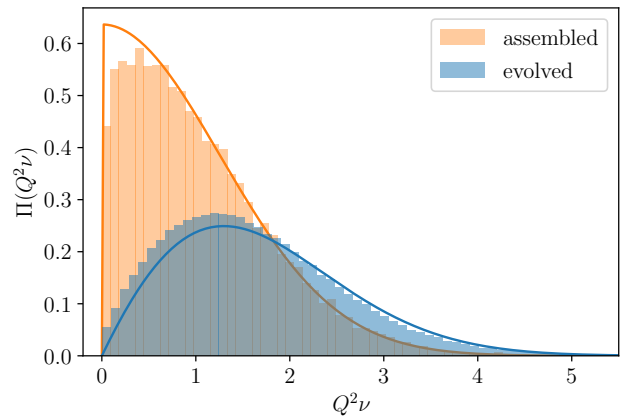

Figure 9. Number density $\Pi(z)$ of the scaled abundance $z = Q^2 \nu$ in $\gamma = \Sigma = 0$ LV model with $Q = 20$: for saturated assembled community (yellow) and evolved Red Queen steady state (blue), with normalization $\int dz \Pi(z) = \overline{L}/Q^2$. Histograms are numerical data and solid lines predictions from, respectively, assembled cavity calculation (Appendix F) and evolutionary cavity calculation (Equation I13). Parameters $Q = 20$, $\gamma = 0$, $\Sigma = 0$.

However, there is a limit of the linearized model that is simple: if we take $D \to \infty$ and $\kappa \to 0$ while keeping the product $\kappa^2 D$ finite, we can rescale the interaction matrix by a factor of $\sqrt{D}$ to obtain the LV model with $\gamma = 0$ and $Q = \kappa\sqrt{D}$, implying $\overline{L} = \mathcal{O}(Q^2)$ in the Red Queen phase. (Note that in this limit, the elements of V become gaussian and the off-diagonal elements are $\mathcal{O}(1)$.) The fact that for $\gamma = 0$ the interaction matrix V has no correlations across the diagonal makes the feedback effects smaller by a factor of $1/\sqrt{L}$ (and of random sign) and hence negligible—therefore the analysis that we undertake becomes tractable. We first treat the case of $\gamma = 0$, $\Sigma = 0$, and then set up the general structure of the cavity calculation for nonzero γ and Σ, before heuristically analyzing the effects of $\Sigma > 0$.

Our analysis of the Red Queen phase proceeds by means of a dynamical cavity method, which is one way of carrying out dynamical mean-field theory (DMFT) as used in systems with quenched randomness [49]. The idea is to consider the evolutionary-time dynamics of a focal "cavity strain" labelled 0, that could potentially enter the community, persist for some number of epochs, and later go extinct. For $\gamma = 0$, the abundance of the cavity strain while it is in the community is determined by its bias, if positive. In this simple case the strain's feedback on itself averages to zero so that $\nu_0 = \xi_0/Q$ in the limit of large L (Appendix F). However the bias, $\xi_0(T) = s_0 + \zeta_0(T) - \Upsilon(T)$, changes in evolutionary time T. If the strain attempts to invade at epoch $\mathbb{T}_0$, then its abundance in the community at all later times is

$$\nu_0(T) = \frac{\xi_0(T)}{Q} \mathbb{1}\left[\xi_0(T') > 0 \ \forall \ T' \in (\mathbb{T}_0, T)\right] \qquad (9)$$

where $\mathbb{1}[E]$ is the indicator function which is unity when its argument is true and zero otherwise. The cavity strain's abundance is only nonzero as long as its bias remains positive—and it goes irrevocably extinct once its bias goes negative. Therefore the statistics of the cavity strain abundance are determined by the statistics of its bias together
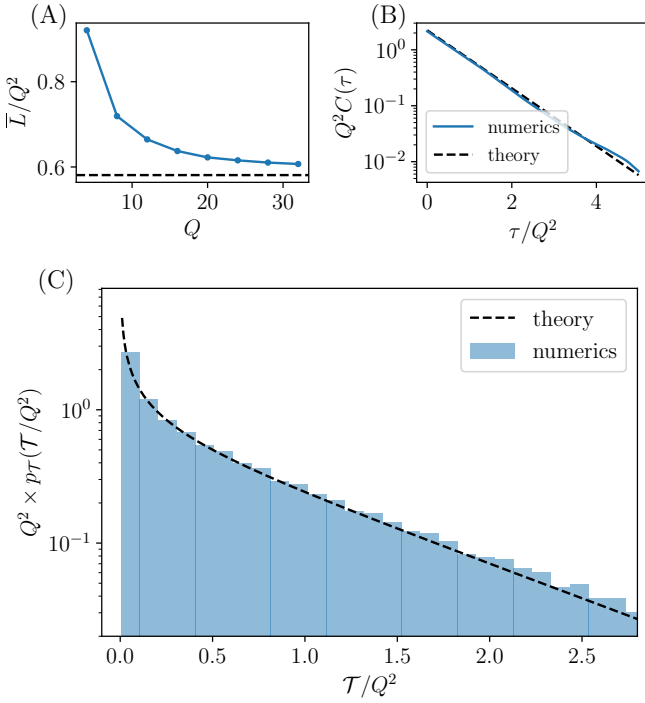
Figure 10. Predictions of evolutionary cavity calculation and simulations for Red Queen state of $\gamma = 0$, $\Sigma = 0$, LV model with independent invaders. Niche parameter that sets time and diversity scales is $Q = 20$ except in (A): Number of strains as function of Q converges to predicted $\overline{L} \approx 0.58 \times Q^2$ (dotted black line) as $Q \to \infty$, Equation 13. (B) Temporal correlation function $C(\tau) = \sum_j \nu_j(T)\nu_j(T + \tau)$ agrees with Equation 12 with no fit parameters. (C) Distribution of strain lifetimes $\mathcal{T}$, normalized by basic timescale Q^2, agrees with prediction of Equation 14.

with the permanent extinction condition.

To carry out the cavity analysis, one needs to make an *Ansatz* for the statistics of the cavity bias, and then explicitly calculate the statistics of the cavity strain's abundance, filtered through the extinction condition. One then enforces self consistency, which relates the extinction-conditioned dynamics of strain abundances in the community to the original statistics of the cavity bias. This is possible since the cavity drive is simply $\zeta_0(T) = \sum_{j \in \mathcal{C}(T)} V_{0j}\nu_j(T)$, which has gaussian correlations for large L. Since the V_{0j} are independent and not conditioned on properties of the community, we have

$$C(\tau) \equiv \overline{\zeta_0(T)\zeta_0(T+\tau)} = \sum_j \overline{\nu_j(T)\nu_j(T+\tau)} \qquad (10)$$

where the sum is over *all* strains that have invaded before time T—most of which will be extinct at T or $T + \tau$ and therefore not contribute to the correlation function. For large $\overline{L}$, the average over T, denoted by the overbar, is not needed: we expect $C(\tau)$ to be "self-averaging" and depend only on the time-difference τ (up to subdominant corrections).

The self-consistency condition at the heart of the cavity method is that the autocorrelation of the cavity strain, averaged over trajectories $\{\zeta_0(T)\}$ and invasion time $\mathbb{T}_0$, is the same as the autocorrelation function of a strain i with

$s_i = s_0$, that determines $C(\tau)$ via Equation 10. $\overline{\Upsilon}$ has to be adjusted to fix $\sum_i \nu_i = 1$ (just as Υ is determined for the static ecological cavity analysis of Appendix F).

We now specialize to the simplest case where all s_i are zero. Simulations of the $\gamma = 0$, $\Sigma = 0$ Lotka-Volterra model (Figure 15B) suggest that in this case $\overline{\Upsilon} = 0$ and we make the *Ansatz* that this is exact. Hoping for some luck, we then make the simplest possible *Ansatz* for the statistics of $\zeta_0(T)$: namely that it has an exponentially decaying correlation function $C(\tau) = \sqrt{J/b} \times e^{-b\tau}$. This *Ansatz* for the correlation function can be translated into a stochastic process that determines ζ_0:

$$\frac{d\zeta_0}{dT} = -b\zeta_0 + \sqrt{2J}\eta_0(T) \qquad (11)$$

with T the time measured in epochs and $\eta_0(T)$ white noise satisfying $\langle \eta_0(T)\eta_0(T') \rangle = \delta(T - T')$ (angular brackets here denote averages over the stochasticity caused by invaders). The white noise represents the changes in the community from adding independent invaders which cause the diffusive changes in abundances discussed in Section IV B. The deterministic term is, roughly, regression to the mean ζ_0 of zero, which can be seen to arise from the correlation between ζ_0 and $\delta\zeta_0$ due to the same elements of V appearing in both quantities [43]. The quantities b and J are (for now) phenomenological parameters that must be fixed later by self-consistency. The advantage of having a Markovian stochastic differential equation for ζ_0, along with $\overline{\Upsilon} = 0$, is that the method of images can be used to find the statistics of ν_0 with an absorbing boundary condition at zero, conditioned on its invasion at time $\mathbb{T}_0$.

In order to calculate the right hand side of the self-consistency condition Equation 10, we need to calculate the two-time correlation function of ν_0. Since all strains are statistically equivalent, we can calculate $\langle \nu_0(T)\nu_0(T + \tau) \rangle$ for the cavity strain entering the community with the initial condition of $\zeta_0(\mathbb{T}_0)$, averaged over the stochasticity in Equation 11. We then average this correlation function over $\mathbb{T}_0$ and $\nu_0(\mathbb{T}_0)$. We know that $\nu_0(T)$ is determined by Equation 9 with $\xi_0(T) = \zeta_0(T)$ since $s = \overline{\Upsilon} = 0$. From our *Ansatz*, the initial $\nu_0(\mathbb{T}_0)$ is distributed as a gaussian random variable with standard deviation $\frac{1}{Q}\sqrt{J/b}$, truncated to be positive (as epochs are measured per successful invasion). In order to calculate $C(\tau)$ from the strains in the community, we integrate $\langle \nu_0(T)\nu_0(T + \tau) \rangle$ over all initial times $\mathbb{T}_0$—equivalent to fixing $\mathbb{T}_0 = 0$ and integrating over all T—and over all starting abundances $\nu_0(\mathbb{T}_0)$ weighted by the truncated gaussian measure. We work with all times in evolutionary epochs, so that the integral of the distribution of $\nu_0(\mathbb{T}_0)$ for $\nu_0(\mathbb{T}) > 0$ is the rate of successful invasions per unit time, which is 1 in our units.

A self-consistent solution is possible if the result of this calculation gives an exponentially decaying correlation function that can be equated with the correlation function of the cavity drive $C(\tau) = \sqrt{J/b}\ e^{-b\tau}$. Fortunately, this turns out to be the case. An additional self-consistency condition fixes $\sum_i \nu_i = 1$. Put together, these conditions allow us to solve for J and b. We then compute the diversity $\overline{L} = \sum_i \Theta(\nu_i)$. These calculations are carried out in Appendix I. The key result is the self-consistent exponential

solution for the correlation function

$$C(\tau) = \frac{2\pi\lambda^2}{Q^2} \exp\left[-\frac{2\lambda}{Q^2}\tau\right], \tag{12}$$

with $\lambda = \frac{1+2\log 2}{4}$, and that the steady state diversity is

$$\overline{L} = \frac{2\log 2}{1+2\log 2}Q^2 \approx 0.58 \times Q^2. \tag{13}$$

The timescale of the decay of the correlation function is $\mathcal{O}(Q^2) = \mathcal{O}(\overline{L})$, consistent with the community turning over every $\mathcal{O}(\overline{L})$ successful invasions. We additionally find that $b = 2\lambda/Q^2$ and $J = 4\pi\lambda^3/Q^4$.

Figure 10 compares the analytical results to numerics on the $\gamma = 0$, $\Sigma = 0$ LV model. As Q increases in Figure 10A, the numerics get closer to the large-$\overline{L}$ regime in which the theory holds, and $\overline{L}/Q^2$ in steady state approaches $\frac{2\log 2}{1+2\log 2}$ from above. Figure 10B shows the scaled correlation function $Q^2 C(\tau)$ versus scaled time τ/Q^2, which agrees well with Equation 12.

From the self consistent solution we can also calculate the distribution of strain *lifetimes*, where the lifetime of a strain i, denoted $\mathcal{T}_i$, is the number of evolutionary epochs for which it has nonzero abundance. In Appendix I we find a distribution

$$p_{\mathcal{T}}(\mathcal{T}) = \frac{2b}{\pi\sqrt{e^{2b\mathcal{T}} - 1}}. \tag{14}$$

Figure 10C shows good agreement between this analytical result and the numerics: $p_{\mathcal{T}}(\mathcal{T})$ decays exponentially at large $\mathcal{T}$ as expected, and has an integrable divergence as $\mathcal{T} \to 0$, due to the distribution of invaders having a constant density at low abundance. These invaders go extinct very quickly as discussed in Section IV B. Note that, as successful invaders must balance extinctions, $\langle\mathcal{T}\rangle = \frac{\log 2}{b} = \overline{L}$: the mean diversity is equal to the average strain lifetime in epochs.

The dynamic cavity method also enables us to determine the distribution of strain abundances in steady state, as the solution to a Fokker Planck equation with appropriate boundary conditions and with a source term for the invading strains (Equation I12). This solution is plotted in Figure 9, in reasonable agreement with the numerical result for the distribution of abundances in the Red Queen phase. As anticipated in Section IV B, there is a marked difference between the abundance distribution in the eco-evolutionary steady state and the saturated assembled community abundance distribution with $L = \hat{L}$. In the assembled community the abundance distribution is a truncated gaussian while in the Red Queen steady state, the density of strains vanishes at low abundance due to the constant jostling of the community by invaders which knock low-abundance strains extinct: the effective diffusion coefficient capturing this process is proportional to the fragility.

For the LV model with $\gamma = 0$, the fragility can be calculated directly due to the simple relationship between ν_i and ξ_i. For each strain with positive bias, $\nu_i = \xi_i/Q$ so the susceptibility to a change in growth rate is $1/Q$ (Appendix F). However the change in each ν_i causes a small random change of the biases of all the other strains, which increases the effective variance of the random field, and

so on. Summing up these effects [43] yields a total effect of random h_j's given by $\sum_i (\delta\nu_i)^2 = \Xi\,\mathrm{Var}[h_j]$ with the fragility $\Xi = \frac{L/Q^2}{1-L/Q^2} = \frac{L}{Q^2-L}$. This result does not depend on how the community is formed, or on the selective differences if they are present. It shows right away the maximum size of a stable community is $\hat{L} = Q^2$, which is saturated for an assembled community. For the evolved community, Ξ cannot be large—otherwise multiple strains would be driven extinct by each invader. Thus $\overline{L} < \hat{L}$, as we find in numerics.

D. General cavity formulation

We now generalize the cavity analysis away from $\gamma = \Sigma = 0$. For $\gamma \neq 0$ in the LV model the analysis is particularly complicated, and for $\kappa > 0$ in the linearized model even more so as both resources and strain dynamics have to be analyzed, as in the ecological cavity calculation (Appendix F). The bias and abundance of a focal strain are no longer simply related as in the $\gamma = 0$ case: due to the feedback of the strain on itself via its influence on the other strains, $\nu_0(T)$ will depend both on its bias $\xi_0(T)$ and an additive term arising from the feedback integrated over all times since its invasion. Therefore the analog of Equation 9 for $\gamma \neq 0$ is

$$\nu_0(T) = \frac{\phi_0(T)}{Q}\mathbb{1}\left[\phi_0(T') > 0\ \forall\ T' \in (\mathbb{T}_0, T)\right] \tag{15}$$

where

$$\phi_0(T) = \xi_0(T) + \gamma\int_{\mathbb{T}_0}^{T} dT'\, M(T-T')\nu_0(T') \tag{16}$$

with $M(\tau)$ a memory kernel that must be determined self-consistently. It will include an instantaneous—on evolutionary time scales—part from the invader's *ecological* feedback on its own abundance, in addition to a slow memory part that extends over many epochs [17]. Both the memory effects and the $s_0 - \Upsilon$ part of ξ_0 mean that the cavity strain will not go extinct when ζ_0 goes through zero, but instead at some other time when $\phi_0(T)$ first goes negative.

The statistics of the drives $\zeta_0(T)$ are determined by the extant community, but $C(\tau)$ is no longer a simple exponential. Nevertheless, one can write an effective stochastic equation for ζ_0:

$$\frac{d\zeta_0}{dT} = -\int_{\mathbb{T}_0}^{T} dT'\, B(T-T')\zeta_0(T') + \sqrt{2J}\eta_0(T) \tag{17}$$

with $B(\tau)$ another memory kernel and $\eta_0(T)$ still white noise which represents the diffusion-like effects of an invader on the other strains discussed in Section IV B. The Laplace transform of B is related, via a Weiner-Hopf factorization, to the Fourier transform of C.

In principle, to proceed one assumes a $C(\tau)$—or equivalently a $B(\tau)$ and J—as well as an $M(\tau)$ and Υ. Then one computes $\langle\nu_0(T)\nu_0(T+\tau)\rangle$ conditioned on ν_0 invading at time $\mathbb{T}_0$. By averaging over $\nu_0(\mathbb{T}_0)$, $\mathbb{T}_0$ and s_0, one can enforce the self-consistency of $C(\tau)$ using Equation 10, as

for the $\gamma = 0$ case. Then $M(\tau)$ is self-consisted from calculating the response of ν_0 to a change in its bias, and Υ is self-consisted to ensure $\sum_i \nu_i = 1$. In the linearized model with resources, the resources have to be handled by a cavity method as well but their effects can be integrated out to give a more complicated relation between strain abundances and M, with Q replaced by $\kappa\sqrt{D}$ and γ replaced by κ^2 (Appendix F).

Unfortunately, even numerical analysis of the self-consistency conditions to determine the memory functions M and B is likely challenging [49]. However the above structure allows us to glean some insights.

In the absence of selective differences, all strains are statistically similar and the systematic decrease in ζ_0 should make the strain lifetime distribution decay exponentially at long times. This suggests that both the memory kernels M and B will decay exponentially at long times, as will the correlation function of the drives, $C(\tau)$. To explore the effects of feedback, we carried out simulations of the $\Sigma = 0$ LV model for $\gamma = -0.8$, 0, and 0.5, and plotted $C(\tau)$ in Figure 19B. We see that $C(\tau)$ indeed decays exponentially at long times for all these γ, with some deviations at shorter times in opposite directions for positive and negative γ. Concomitantly, as γ increases, the timescale over which $C(\tau)$ decays, scaled by the overall turnover time of $\overline{L}$ epochs, grows.

There is no sign of marginality in the exact solution for $\gamma = \Sigma = 0$—therefore the Red Queen phase should be robust to small perturbations, as evinced by these simulations for $\gamma \neq 0$. The heuristic arguments below suggest that the robustness extends as well to arbitrarily large Σ, and numerical results suggest that it extends also to all $\kappa < 1$ for the linearized model and up to modest—although *not all*—γ in the LV model: see Section V for more discussion of this point.

E. Effects of selective differences

The Red Queen phase that obtains for $\Sigma > 0$ in both the linearized and LV models has many properties in common with the $\Sigma = 0$ case we have analyzed. Here we discuss more systematically the effects of a distribution of general fitnesses, focusing first on the correlation function $C(\tau)$, whose decay at large τ is dominated by strains with anomalously large s. We then consider large Σ and give scaling arguments for how the community evolves into the tail of the s distribution which allows us to find the dependence of $\overline{\Upsilon}$ and $\overline{\log p_{inv}}$ on Σ. All of these steady-state properties—including the decay of $C(\tau)$—are determined by the behavior of the $p(s)$ distribution at large s—thus we also consider non-gaussian $p(s)$.

Correlation function for $\Sigma > 0$: From the general structure of the cavity analysis, we can understand the behavior of a strain with anomalously large s (we suppress the strain index for notational convenience). The behavior is simplest for $\gamma = 0$ for which the only memory is in $\zeta(T)$, which still has mean zero. Extinction occurs when $\xi \searrow 0$, i.e. $\zeta \searrow \Upsilon - s$ which is unlikely for large s: for $s - \Upsilon \gg \mathrm{Std}[\zeta]$, ξ will fluctuate around $s - \Upsilon$, staying far from zero for a long time. The rare event for ζ to decrease to $\Upsilon - s$ has, quite

generally for a stationary gaussian process, probability of order $\sim e^{-k(s-\Upsilon)^2/\mathrm{Var}[\zeta]}$ with the coefficient k depending on the statistics of ζ governed by $B(\tau)$. The correlation function of a strain with general fitness s_i therefore decays exponentially with a timescale that is gaussian-long in s_i. With a distribution of general fitnesses that extends out to large s, the average of $C(\tau)$ over all the strains is dominated at long τ by the fraction of strains whose lifetime is longer than τ. For gaussian $p(s)$ the long-lived strains are gaussian-rare in s and have a lifetime that is gaussian-long in s: averaging over $p(s)$ then yields a $C(\tau)$ that decays as a Σ-dependent power law for large τ, as shown in Appendix K. Note that for an exponential $p(s)$, which will play a role in the analysis for large Σ, one finds that $C(\tau) \sim e^{-a\sqrt{\log \tau}}$ with a depending on Σ.

Associated with the long tail of the correlation functions will be long tails to the memory kernels. Thus for γ or κ non-zero and $\Sigma > 0$, it is not obvious whether or not the persistent memory in $M(\tau)$ will change the behavior qualitatively. To explore the effects of increasing Σ when $\kappa > 0$, in Figure 19A $C(\tau)$ is plotted with gaussian $p(s)$ for various Σ in the linearized model with $\kappa = 0.8$: one sees a slower decay of $C(\tau)$ for increasing Σ. Although this is far from conclusive, we conjecture that the feedback-induced memory effects do not fundamentally change the effects of anomalously large s's in the communities.

The analysis above suggests that the mean strain lifetime, $\langle \mathcal{T} \rangle$, becomes infinite for sufficiently wide gaussian $p(s)$, and certainly for exponential $p(s)$. This regime must involve additional subtleties because the naive divergence of $\langle \mathcal{T} \rangle$ is inconsistent with the expected relation $\overline{L} = \langle \mathcal{T} \rangle$. One way this could happen might be if the long-tailed memory increased the rate of extinction through anomalously long-lasting negative ζ_i which could drive anomalously large $s_i - \Upsilon$ extinct at a higher rate than the naive "barrier-crossing" estimate. Another possibility is that the large D limit is rather more subtle due to effects of rare large s_i's. However, we have not explored these issues in depth.

Evolutionary transients and steady state for $\Sigma > 0$: We now analyze the evolutionary behavior in the limit of large Σ. In what follows we will mainly discuss the linearized model in which the distribution of biases has $\mathcal{O}(\sqrt{D/L}) = \mathcal{O}(1)$ width—this can be translated to the LV model by recalling that there the width of the bias distribution is $\mathcal{O}(1/Q)$.

For $\Sigma \gg 1$, there is a long evolutionary transient during which the distribution of extant s_i creeps into the tail of $p(s)$. The tail of $p(s)$ therefore determines the long term evolutionary dynamics, including—when it occurs—the properties of the Red Queen phase. As the population-mean s (and therefore Υ) increases, the extant s's will be near to Υ with the scale of their variation set by the shape of $p(s)$ near Υ. The width of the extant s_i distribution will be of order

$$\Delta(\Upsilon) \equiv \mathrm{E}[s_I - \Upsilon | s_I > \Upsilon] \approx -\left[\left. \frac{d\log p(s)}{ds} \right|_{s=\Upsilon} \right]^{-1}, \quad (18)$$

which decreases with increasing Υ for $p(s)$ decaying faster than exponentially and increases with Υ for $p(s)$ decaying slower than exponentially. At any point in the evolution,

once $\Upsilon \gg \Sigma$, the distribution of extant s_i is concentrated around Υ and the system behaves as if there were an exponential distribution of the s_i with scale $\Delta(\Upsilon)$.

Given this exponential approximation to the general fitness distribution, it is helpful to understand the behavior of an *assembled* community with an exponential $p(s)$ when the extant s_i are large enough that the lower bound of $p(s)$ does not matter—this will determine a key relationship between L and Υ. If we assemble a large number of strains with an exponential distribution of general fitnesses $p(s) = \frac{1}{\Delta} e^{-s/\Delta} \Theta[s]$, for sufficiently large Δ the ecological cavity analysis (see Appendix G) yields an assembled community with $L = \kappa D/\Delta$ in the linearized model (and $L = Q/\Delta$ in the LV model). We therefore expect that during the transient evolution, the community will reach a slowly-changing quasi steady-state with diversity given, at any point in time, by $\kappa D/\Delta(\Upsilon)$.

We have shown that the behavior of Υ determines the behavior of L. How does Υ depend on T during this transient and into the steady state of the evolution? The increase of Υ is tied to the increasing mean s, denoted $\langle s \rangle \equiv \sum_i \nu_i s_i$, so we are led to analyze how $\langle s \rangle$ changes in evolutionary time. Each successful invader I must have $\zeta_I + s_I > \Upsilon$. The probability density of s_I conditioned on invasion is (see Appendix of ref. [43])

$$p(s_I | \text{invasion}) \approx \frac{1}{\Delta} \exp\left[-\frac{s_I - \Upsilon}{\Delta} - \frac{\sigma_\zeta^2}{2\Delta^2} \right] \Phi\left[\frac{s_I - \Upsilon}{\sigma_\zeta} \right], \quad (19)$$

where $\sigma_\zeta^2 \equiv \text{Var}[\zeta_I] = \sum_i \nu_i^2 \approx D/L$ is the Simpson index of the community and Φ is the CDF of the standard normal distribution. It follows that

$$E[s_I - \Upsilon | \text{invasion}] = \Delta(\Upsilon) - \frac{\sigma_\zeta^2}{\Delta(\Upsilon)}. \quad (20)$$

As a result, when $\Delta(\Upsilon) \approx \sigma_\zeta \approx 1/\kappa$, there is a steady state and Υ is roughly constant in evolutionary time, with its value depending on the form of the function $\Delta(\Upsilon)$. During the transient and in steady state, $\overline{L} \sim \kappa D/\overline{\Delta}(\Upsilon) = \mathcal{O}(\kappa^2 D)$.

We can now proceed to analyze the behavior of a family of distributions $p(s) \sim e^{-(s/\Sigma)^\psi/\psi}$, for which $\Delta(\Upsilon) \approx \Sigma^\psi/\Upsilon^{\psi-1}$. For stretched exponential decay of $p(s)$, corresponding to $\psi < 1$, $\Delta(\Upsilon)$ is an increasing function—therefore Υ will run away to positive values with L concomitantly decreasing. The diversity crashes and eventually the community will reduce to a single strain that goes extinct when a fitter strain invades. Even if Σ is small and seemingly unimportant, eventually a rare large s will emerge, leading to Υ large enough that $\Delta(\Upsilon) \gtrsim 1$, and therefore to decreasing diversity.

For $\psi > 1$ (including the gaussian case $\psi = 2$), initially Δ is of order Σ. From the above analysis, we expect diversity $L \sim \kappa D/\Delta$. With Δ large, invasions will tend to increase Υ, thereby decreasing Δ and increasing L. Eventually Δ will decrease to $\overline{\Delta} \sim \sqrt{D/L} \sim 1/\kappa$, independent of ψ, so that the widths of the extant s_i and ζ_i distributions are comparable. Concomitantly, L will saturate at $\overline{L} \sim \kappa D/\overline{\Delta}$, and Υ will saturate at $\overline{\Upsilon} \sim (\Sigma^\psi/\overline{\Delta})^{1/(\psi-1)}$, with $\overline{\log p_{inv}} \sim (\Sigma/\overline{\Delta})^{\psi/(\psi-1)}$.

We are now in a position to analyze the transient evolutionary dynamics that occur when $\Sigma \gg 1$ and $\psi > 1$. For large Δ, as it will be initially, the change in Υ due to a successful invasion by strain I is roughly $\Delta(\Upsilon)/L$, coming from the fitness increment multiplied by the $\mathcal{O}(1/L)$ typical abundance ν_I. Since the fitness increment is $\sim \Delta$ and $1/L \sim \Delta/D$, the change in Υ per invasion is $d\Upsilon/dT \sim \Delta(\Upsilon)^2/D$. On average $d\Upsilon/dT$ is positive until the contribution of ζ_I becomes comparable to $s_I - \Upsilon$, at which point $\Upsilon \approx \overline{\Upsilon}$ and $\Delta \approx \overline{\Delta}$.

The transient behavior is found by integrating up the typical change per invasion: roughly $d\Upsilon/dT \sim \Delta(\Upsilon)^2/D \sim \Sigma^{2\psi}/D\Upsilon^{2\psi-2}$ until Υ is close to $\overline{\Upsilon}$. Thus during the transient

$$\Upsilon \sim \left(\frac{\Sigma^{2\psi} T}{D} \right)^{\frac{1}{2\psi-1}} \quad (21)$$

$$\frac{L}{D} \sim \Sigma^{-\frac{\psi}{2\psi-1}} \left(\frac{T}{D} \right)^{\frac{\psi-1}{2\psi-1}} \quad (22)$$

for a duration

$$T_{transient} \sim D\Sigma^{\psi/(\psi-1)}, \quad (23)$$

after which the steady state ensues.

An important observation is that some aspects of the Red Queen steady state for large Σ are *independent of details* of the s distribution and of Σ: in particular, for large Σ, $\overline{L}$ is independent of Σ, as seen in Figure 4 for gaussian $p(s)$, since $\overline{\Upsilon}$ becomes large enough that the distribution of the extant s_i has $\mathcal{O}(1)$ width, irrespective of Σ.

To compare with our numerics for gaussian $p(s)$, we specialize to $\psi = 2$ in which case $\Delta(\Upsilon) \approx \Sigma^2/\Upsilon$. Here we have $\overline{\Upsilon} \sim \Sigma^2$ and $\overline{p_{inv}} \sim e^{-q\Sigma^2/2}$ for q depending on κ. This scaling is similar to the saturated assembled community (Section IV A), in which $-\log \hat{S} \sim \Sigma^2$ was conjectured to scale with $\overline{\log p_{inv}}$. During the long initial transient, $\Upsilon \sim T^{1/3}$ and $L \sim T^{1/3}$ with the transient lasting for a number of epochs $T_{transient} \sim D\Sigma^2$, consistent with the rapid growth of transient length with Σ in Figures 4, 5 and 6. In Appendix J, we check that this predicted transient behavior of Υ agrees with numerics.

Because of large κ-dependent coefficients, even moderate values of Σ can make p_{inv} unrealistically small at steady state. For $\Sigma = 1$, $\rho = 0$ and $\kappa = 0.8$, the invasion probability at steady state is less than 10^{-15} (Figure 5D) which suggests that $\overline{\Delta}$, which determines $\overline{\Upsilon}$ via $\overline{\Upsilon} \sim \Sigma^2/\overline{\Delta}$, has a value $\overline{\Delta} \cong 0.12$ for $\kappa = 0.8$, substantially smaller than the value for the top-down assembled community, $\overline{\Delta} = (1 - \kappa^2)/\kappa = 0.45$ (Equation 8a).

In the CR model, invasion probabilities are even smaller: as discussed earlier, the width of the general fitness terms dominates over that of the interactions by a factor of $\sqrt{D}$, corresponding to $\Sigma \sim \sqrt{D}$ in the linearized model. Therefore, for our gaussian ensemble, $\overline{p_{inv}}$ is exponentially small in D. Indeed we see (Figure 2B) that for $\kappa = 0.4$ and $D = 50$ in the CR model, the invasion probability in steady state is $\sim 10^{-400}$, where here the distribution of general fitnesses U_i inherits its gaussian distribution from the gaussian ensemble of the $\mathcal{G}$ matrix that we studied.

Generally, distributions with faster-decaying tails but similar Σ have higher invasion probability at steady state.

However even in the limit $\psi \to \infty$ which corresponds to a double exponential decay $p(s) \sim \exp(-\exp(s/\Sigma))$, $\overline{p_{inv}}$ still decays exponentially in Σ, and therefore as a stretched exponential in D for the CR model.

For distributions that have a sharp cutoff at some maximum s, the invasion probability in steady state is not so small, even with a broad distribution of s's. If $p(s) \sim (\Sigma - s)^\lambda \Theta(s)\Theta(\Sigma - s)$ for large Σ and some power λ, then the width of the extant s_i distribution scales as $\frac{1}{\lambda+2}(\Sigma-\Upsilon)$. The Red Queen phase obtains when the distribution of extant s has $\mathcal{O}(1)$ width, in which case $p_{inv} \sim 1/\Sigma^{\lambda+1}$, which decays with Σ much more slowly than in the case of $p(s) \sim e^{-(s/\Sigma)^\psi/\psi}$. Importantly, once the Red Queen steady state is established, the population-mean $\langle s \rangle$ does not creep ever closer to the cutoff at the maximum s. This is in contrast to the situation with $\kappa = 1$, where Υ would inch up as $\langle s \rangle$ increases towards the maximum s, since the Lyapunov function Υ has a part that is exactly $\langle s \rangle$.

We have found that there is a wide range of scalings of $\overline{\Upsilon}$ and $\overline{p_{inv}}$ with Σ in the linearized model or with D in the CR model. These scalings can be reasonably obtained with different distributions of the s_i or the elements of $\mathcal{G}$ and $\mathcal{F}$. We have used gaussian distributions since they allow us to easily draw invaders conditioned on having positive invasion fitness, and to explore issues of *principle* regarding the Red Queen phase. However it is important to note that quantitative aspects such as the scaling of the invasion probability with D are very non-universal. In biological systems, situations where the evolution pushes far into the tail of $p(s)$ are anyway likely dominated by effects that we do not account for, such as the discreteness of genomes.

V. DISCUSSION

In this work we have explored evolution in a range of ecological models which exhibit diverse stable communities, and found that, generically, after a transient period of evolution, the community converges to a Red Queen state of constant turnover with the probability of new strains successfully invading roughly constant. In contrast to previous work which finds continual evolution in selection-driven host-pathogen models [26, 52, 53], the diversity we find here stably coexists and does not rely on continual mutations to replenish it. There is no clear notion of kill-the-winner dynamics in which one subpopulation "chases" the other through trait space; nonetheless constant turnover ensues. The evolving ecology acts as a concrete manifestation of the environmental feedback that can engender continual Red Queen evolution in random landscape models [17] without diversity. The consumer resource model and a linearized approximation to it display a Red Queen phase as long as the consumption and growth are not perfectly correlated ($\kappa < 1$): the perfectly correlated case often studied behaves differently, not exhibiting a Red Queen phase. The Lotka-Volterra model with random interactions also exhibits a Red Queen phase but only for symmetry parameter less than a critical value.

Importantly, the Red Queen phase persists in the presence of arbitrarily large general fitness differences s_i provided that their distribution falls off faster than exponen-
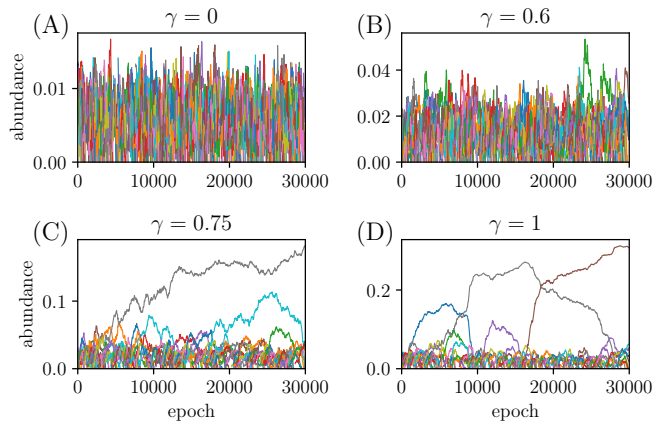

Figure 11. Strain trajectories (subset of strains shown) for LV model with various γ in Red Queen (A)–(B) and oligarch (C)–(D) phases. In oligarch phase, an $\mathcal{O}(1)$ fraction of the total abundance "condenses" into a few strains which continue to turn over more and more slowly. $Q = 20$ and $\Sigma = 0$ throughout.

tially for large s. Previous analyses of high-dimensional ecological models [19, 36] have, as noted earlier, first assumed that in the large D (or similar) limit, the distribution of general fitnesses (or similar quantities) has a width that is smaller than that of the interactions by a factor of $1/\sqrt{D}$, and then taken a well-behaved large-D limit. We have shown that this assumption is not needed: In a long-evolved community, the width of the distribution of general fitness differences *within the evolved community* will become comparable to the variations between strains of the effects of ecological interactions. Even starting with a parametrically (in D) broad distribution of general fitnesses, some properties (such as the diversity) of evolved communities are independent of the distribution, though other properties (such as the invasion probability of new strains) can be strongly affected.

If the distribution of the s_i falls off with a large characteristic scale—which occurs naturally in the CR model—there is a long transient period during which evolution is dominated by increases in $\langle s \rangle$ and the diversity gradually increases as the population moves out into the tail of $p(s)$. The transient behavior converges to a steady state when the strain-to-strain variations of s_i and the ecological influences of other strains in the community are comparable—here the diversity of the Red Queen steady state becomes independent of $p(s)$, but the invasion probability at steady state can be extremely small. However if there is a maximum to s, the invasion probability, and hence the community turnover rate, stabilizes at a value that need not be as small.

Below, we discuss some unresolved issues from this current work, and suggest future directions.

Other eco-evolutionary phases in well-mixed systems: What other long term states occur naturally and robustly in eco-evolutionary models without any spatial structure or externally varying environment? For models with perfect symmetry, such as the $\kappa = 1$ CR and linearized models, the presence of a Lyapunov function means that the community can continue to evolve and turn over but will do so more and more slowly. Does such slowing evolution occur away
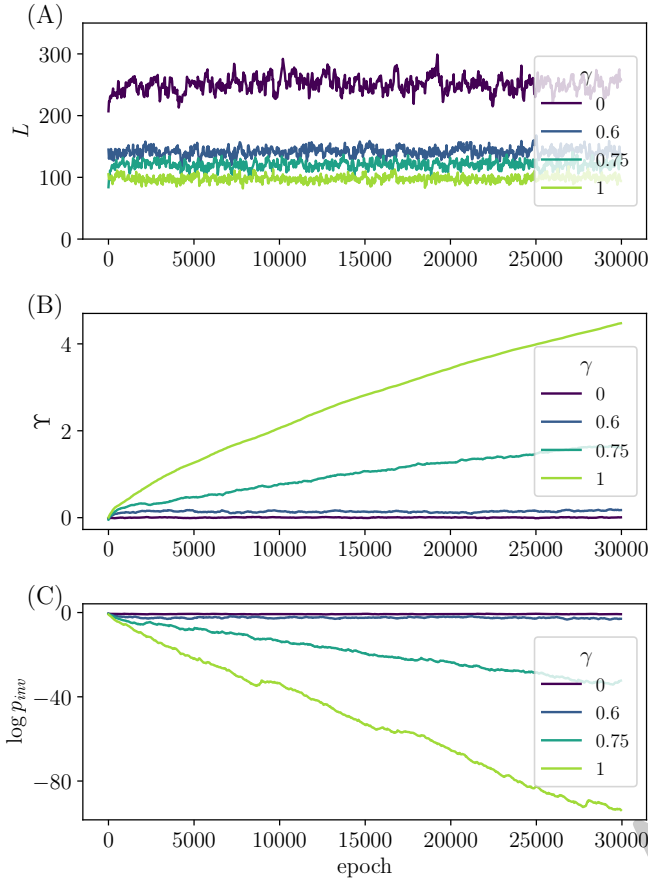
Figure 12. Evolution in LV model ($Q = 20$, $\Sigma = 0$) undergoes a transition from Red Queen for $\gamma < \gamma_c \cong 0.65$ to oligarch phase for $\gamma > \gamma_c$. (A) L is roughly constant in both phases. (B) Υ fluctuates around steady state in Red Queen phase but increases without bound in oligarch phase (C) p_{inv} achieves a steady state in Red Queen phase but is ever-decreasing in oligarch phase.

from special slices of parameter space? In the simple CR models with a convex Lyapunov function, the Red Queen phase appears to replace such a state as soon as $\kappa < 1$. However for LV models, we have found evidence for sharp (in the limit of large Q) qualitative change as a function of γ, around $\gamma_c \approx 0.65$ for $\Sigma = 0$. For $\gamma < \gamma_c$, the Red Queen phase obtains and there is continual turnover with the mean general fitness, Υ, and the invasion probability, p_{inv}, roughly constant. However for $\gamma > \gamma_c$, we find no steady state. This transition is not apparent in the dynamics of the diversity, which is still roughly steady over time, but shows up strikingly in other properties.

For $\gamma > \gamma_c$, the dynamics continually slows down, with Υ ever-increasing and p_{inv} thereby decreasing so that the community becomes progressively harder to invade. This behavior, which we have dubbed an *oligarch* phase, has surprising features that we do not yet understand. Most strikingly, the evolutionary dynamics show that a few strains— the oligarchs—become far more abundant than the others.

Figure 11 shows the abundance dynamics in both the Red Queen and oligarch phases, and Figure 12 the corresponding trajectories of L, Υ and p_{inv}. For $\gamma = 0$ and 0.6, the constant turnover of the Red Queen phase is evident in the narrow distribution of strain abundances and lifetimes.

However for $\gamma = 0.75$ and 1, an $\mathcal{O}(1)$ fraction of abundance "condenses" into an $\mathcal{O}(1)$ number of oligarch strains which still slowly turn over.

Qualitatively, the oligarch phase appears to occur due to feedback between the abundant strains and the invaders for $\gamma \to 1$: if a prospective invader, I, is to successfully enter the community it must have a positive contribution to its invasion fitness from an abundant strain ℓ, i.e. $V_{I\ell} > 0$. Then since γ is strongly positive, the invader will tend to have $V_{\ell I} > 0$ as well, thus increasing the abundance of the oligarch strain ℓ. However this argument does not explain the turnover of oligarchs, nor their absence in the linearized model. The non-steady-state nature of the oligarch phase makes analysis difficult since correlation and response functions of strain abundances in evolutionary time depend on two times, not just time differences.

The transition between the Red Queen and oligarch phases (discussed in Appendix E) appears to be first order-like (and, as thus expected, it persists for at least small general fitness variations with γ_c depending on the width, Σ, of their distribution in a way that we have not carefully investigated.) Hysteresis near the transition is observed if one dynamically changes γ through γ_c (Figure 16). Strikingly, if one evolves with $\gamma > \gamma_c$ for some time, and then rapidly quenches to $\gamma < \gamma_c$, one can obtain a steady state which exhibits aspects of both the Red Queen and oligarch phases. A handful of high-abundance strains persists at roughly constant abundance without turning over, while the majority of strains turn over steadily as they would in the "normal" Red Queen phase. In this coexisting state between the two phases, Υ and p_{inv} remain respectively elevated and suppressed relative to their values in the Red Queen phase, while L does not show such a signature (Figure 16). Further analysis of such a steady oligarch state could lead to a better understanding of the oligarch phase as well as the transition between the two eco-evolutionary phases.

What classes of models can exhibit an oligarch phase? Naively, the heuristic argument above would suggest an oligarch phase in the linearized model, though such a phase does not seem to exist. This difference in behaviors is likely related to the properties of the *totally symmetric* LV and linearized models: the linearized model always has a convex Lyapunov function whereas the LV Lyapunov function becomes non-convex for sufficiently large S [37]. What other phenotype models— perhaps with more general interaction via chemicals—can display an oligarch phase of eco-evolutionary dynamics?

If some fraction of the D resource-mediating chemicals have negative correlations between the corresponding columns of G and F, it is possible that in the $D \to \infty$ limit, even with nonvanishing correlations between G and F, the diagonal entries of the interaction matrix $V = -GF^{\intercal}$ are the same order as the off-diagonal entries, in which case one might recover an LV model with $\gamma \neq 0$, which could then display an oligarch phase. Regardless, the presence of the oligarch phase in the LV model illustrates the striking difference in phenomenology that can occur in different models, which can exhibit multiple distinct eco-evolutionary phases, each robust over a range of parameters.

Parent-mutant correlations: In this work, we have pri-

marily studied a process that is not strictly speaking evolution: new strains that are introduced are unrelated to extant strains. Nevertheless, this process is crucially different from well-studied models of top-down assembly. In our pseudo-evolutionary process, strain extinctions are permanent, which qualitatively changes the shape of the abundance distribution. In steady state few strains with low abundance exist as these are most easily driven extinct by invasion-induced perturbations.

We have shown that introducing mutant strains related to extant parent strains (with correlation $\rho > 0$) results in very similar behavior if the mutant and parent are not strongly correlated phenotypically. However if the effects of mutations are small—surely relevant biologically [54]—diversification is slower because of the increased importance of general fitness differences. How do the existence and properties of a diverse Red Queen phase depend on the magnitude of the mutational effects as the correlation between mutant and parent phenotypes approaches unity? What kinds of dynamics obtain for a distribution of mutational effect sizes? What is the interplay between general fitnesses and parent-mutant correlations? What types of phylogenies are generated by the Red Queen eco-evolutionary process? Do they resemble any known class of coalescent trees? We will address these issues in a future paper.

Changing environments: The CR and linearized resource models naturally allow study of how evolving communities are affected by *extrinsic changes* in the environment. During the initial transient period of evolution before the Red Queen phase obtains, the community in the CR model adapts to its resource supply by aligning the growth vectors of strains (rows of $\mathcal{G}$) with the resource supply vector $\mathbf{K}$. If the supply vector $\mathbf{K}$ changes extrinsically after this initial period of evolution, the distribution of U_i will broaden and the diversity will crash, only to rebuild as the community readapts to the new resource supply vector. This scenario can be studied in the linearized model if we parameterize the general fitnesses through an interaction with the environment, and define the resource vector as $R_\alpha = K_\alpha - \sum_j F_{j\alpha}\nu_j$, for resource supply $\mathbf{K}$. This parameterization gives rise to selective differences $s_i = \sum_\alpha G_{i\alpha}K_\alpha$: therefore changing $\mathbf{K}$ will affect the s_i. What do the eco-evolutionary dynamics look like for a continually changing supply vector, either driven extrinsically or through some further feedback from the ecological community? Can this externally driven continual evolution be distinguished from the self-driven Red Queen phase?

Clonal interference: In this work we have taken advantage of a separation of timescales between evolution and ecology. However such a separation is not at all guaranteed to exist, particularly in host-pathogen systems [55], in large microbial populations with abundant supply of beneficial mutations [9], or when mutational effects are small enough that ecological timescales are slow. Investigation of the dynamics of ecological models in a "clonal interference" regime, with multiple ecological mutants emerging before any single one of them has reached its fixed point abundance, is an open challenge of definite relevance to biological systems.

Spatial structure: Spatial structure with spatial transport remains an understudied—although crucial—feature of ecological and evolutionary models. Gut and skin microbiota associated with different individuals are realizations of distinct ecological communities coupled by migration, with new mutants arising locally with the potential to spread across many different people [56, 57].

We have found that even without spatial structure, the eco-evolutionary behavior of stable communities is quite rich. However generalizing our analyses to spatial-mean-field-like "island" models with all-to-all migration—roughly like microbes from person-to-person—is a very interesting direction, as are extensions to two- or three- dimensional spatial structure. Can a diverse Red Queen phase exist if there is only migration of strains between identical spatial patches, but no evolution? Or is evolution required to keep generating new types and produce turnover of biodiversity? If the niche parameter (D or Q^2 in our models) is not large enough for all the strains in the system to coexist locally, then different strains could still inhabit different spatial locations. Is this heterogeneity maintained when strains migrate between spatial patches? When evolution continually produces new variants? How much persistent diversity can be maintained, and how does this grow with the spatial size or number of microbiome hosts?

Other spatiotemporal eco-evolutionary phases: The range of models we have studied in this work hints at the generality of the Red Queen phase in ecological models with stable communities. However, with spatial structure there is another ecologically diverse phase known: a state of spatiotemporal ecological chaos in which the spatial structure stabilizes the fluctuating ecological dynamics against global extinctions, leading to an ecological state that is exponentially long-lived in its spatial size [23]. In previous work [43] we have studied gradual evolution and assembly in such a spatiotemporally chaotic population. In the simplest predator-prey-like models that yield this phase, there are no niches created by different resources: the model corresponds to a $Q = 0$ Lotka-Volterra model with negative correlations across the diagonal ($\gamma < 0$). However the dynamical nature of the ecological phase allows one to pack unlimited diversity into a single "niche," even when strains are added gradually through evolution. In the presence of general fitness differences, diversification to arbitrarily large L can still occur, but the evolution continually slows down, with p_{inv} becoming smaller and smaller as the general fitnesses (and hence Υ) continue to increase.

We can understand the tendency of these host-pathogen-like models towards unlimited diversity from the phase diagram of the top-down assembled community. If one starts in the unstable phase, then increasing the diversity does not bring one closer to the phase boundary between stable and unstable phases; in fact it takes one further from this boundary. There is no natural diversity scale (like D in the resource models) and no divergent fragility that would limit the community's diversity. By contrast, for the stable dynamics studied here, the presence of niches sets a scale for the diversity which is $\overline{L} \sim D$ in the CR and linearized models, and $\overline{L} \sim Q^2$ in the LV model. This niche structure qualitatively changes the outcome of the evolutionary dynamics.

Diversity across scales: Experimental data have shown

that strain dynamics can differ based on the resolution at which they are observed, with closely related strains undergoing larger fluctuations than family-level groupings [54, 58]. Further study of ecological dynamics across scales of diversity is needed. How—hopefully already in simple models—does evolution create diversity over many scales, and how does this diversity interact with the ecological dynamics across such scales? Can evolution robustly lead to ecologically chaotic dynamics of micro-diversity but stable dynamics at coarser levels of differentiation?

Prospects: The models and scenarios described in this work provide a simple framework and examples in which ecology produces an evolutionary Red Queen phase. The analytical tractability of the models in the simplest case—and the generality and robustness of the Red Queen phase we have found—gives hope that the study of simple models might advance our understanding of the evolution and maintenance of fine-scale biodiversity in natural microbial communities.

Appendix A: Structure of CR model

Lyapunov function: The Lyapunov function of the CR model which exists when $\mathcal{G} = \mathcal{F}$ ($\kappa = 1$) is

$$\Lambda = -\sum_\alpha \left[K_\alpha \log \frac{K_\alpha}{\mathcal{R}_\alpha} - (K_\alpha - \mathcal{R}_\alpha) \right] \quad \text{(A1a)}$$

$$\text{with } \mathcal{R}_\alpha = \frac{K_\alpha}{\omega + \sum_{j=1}^S \mathcal{F}_{j\alpha} n_j}. \quad \text{(A1b)}$$

Λ is increasing under the ecological dynamics of Equation 1 and hence also under the evolutionary dynamics even with extinctions, as extinct strains do not contribute [38].

Non-dimensionalizing: In order to study the CR model in more detail, it is useful to identify the combinations of parameters that matter for its fixed points. We make the choice $\omega = 0$, which does not qualitatively change the behavior, but simplifies the model. Then we define $\tilde{\mathcal{R}}_\alpha = \mathcal{R}_\alpha \mu_g/d$, $\tilde\sigma_g = \sigma_g/\mu_g$, $\tilde{\mathcal{G}} = \mathcal{G}/\mu_g$, $\tilde\sigma_f = \sigma_f/\mu_f$, $\tilde{\mathcal{F}} = \mathcal{F}/\mu_f$ and $\tilde K_\alpha = \mu_g K_\alpha/d\mu_f$, such that Equation 1 becomes

$$\frac{dn_i}{d(mt)} = n_i \left(\sum_\beta \tilde{\mathcal{G}}_{i\beta} \tilde{\mathcal{R}}_\beta - 1 \right) \quad \text{(A2a)}$$

$$\frac{d\tilde{\mathcal{R}}_\alpha}{d(\mu_f t)} = \tilde K_\alpha - \tilde{\mathcal{R}}_\alpha \sum_j \tilde{\mathcal{F}}_{j\alpha} n_j. \quad \text{(A2b)}$$

Therefore κ, $\tilde\sigma_g$, $\tilde\sigma_f$ and the rescaled supply vector $\tilde{\mathbf{K}}$ are the only quantities that matter for the fixed point abundances. We can set $\mu_f = \mu_g = 1$ and $d = 1$ without loss of generality. The fixed point abundances $\{n_j^*\}$ and $\{\mathcal{R}_\alpha^*\}$ satisfy the equations

$$\sum_\beta \mathcal{G}_{i\beta} \mathcal{R}_\beta^* = 1 \ \ \forall i \in \mathcal{C}; \quad \mathcal{R}_\alpha^* = \frac{K_\alpha}{\sum_j \mathcal{F}_{j\alpha} n_j^*}. \quad \text{(A3)}$$

We do not expect that the variance of the components of $\mathbf{K}$ plays an important role, so we set $K_\alpha = K$. The value of K matters only insofar as it determines the scale of the $\{n_j^*\}$, which can be rescaled to eliminate K. Therefore in our simulations we set $K = 1$.

Appendix B: Parent-mutant correlations in CR model

We can draw mutant phenotypes to be correlated with the phenotype of a parent, which is chosen with probability proportional to its abundance for each epoch. Given the parent phenotype—the vectors $\mathcal{G}_P$ and $\mathcal{F}_P$—we draw a mutant phenotype $\mathcal{G}_M$ and $\mathcal{F}_M$ which has elementwise Pearson correlation ρ with the parent phenotype. We also enforce elementwise correlation κ between $\mathcal{G}_M$ and $\mathcal{F}_M$. In Figure 13 we show numerical results from this process. The trends are similar to what we observe when we vary ρ in the linearized model and, importantly, show that there is still a Red Queen steady state across a range of ρ, with the phase qualitatively similar to the $\rho = 0$ phase studied in depth. Of particular phenomenological interest is the dependence of $\log p_{inv}$ on ρ, which displays a nonmonotonicity, with the initial increase as a function of ρ similar to the one observed in the linearized model (Figure 6D). We deliberately stay away from the limit $\rho \to 1$, which is left for future work as additional subtleties arise.

Appendix C: Relation between CR and linearized models

We here connect the CR model to the linearized model, which we have argued captures the behavior of the CR model and is easier to analyze. The linearization of $\mathcal{R}^*$ as a function of the $\{n_j^*\}$ is justified when the mean of $\sum_j \mathcal{F}_{j\alpha} n_j^*$ over α (weighted by the K_α) is larger than its variations. In Figure 14A we show that this condition is met in the Red Queen steady state for $\kappa = 0.4$. Figure 14B shows that the total consumer population is roughly constant in evolutionary time, which motivates a description in terms of fractional abundances for the linearized model. We carry out an expansion which allows us to separate out general fitness and interaction contributions to each strain's growth rate.

In the Red Queen steady state the means of $\mathcal{F}$ and $\mathcal{G}$ are greater than 1 due to evolutionary conditioning—for example in Figure 3A, the plot for $\kappa = 0.4$ shows that the mean of $\mathcal{G}$ plateaus at $\langle U \rangle/D \approx 2.5$. The mean of $\mathcal{F}$ is smaller than that of $\mathcal{G}$ but still larger than 1. If we define $\langle \mathcal{F} \rangle \equiv \frac{1}{NK_{\text{tot}}} \sum_{j,\beta} \mathcal{F}_{j\beta} n_j^* K_\beta$ and $f_{i\alpha} = \mathcal{F}_{i\alpha} - \langle \mathcal{F} \rangle$, then we have $\sum_j \mathcal{F}_{j\alpha} n_j^* = \langle \mathcal{F} \rangle N + \sum_j f_{j\alpha} n_j^*$. Defining $\mathcal{N} = \langle \mathcal{F} \rangle N$,
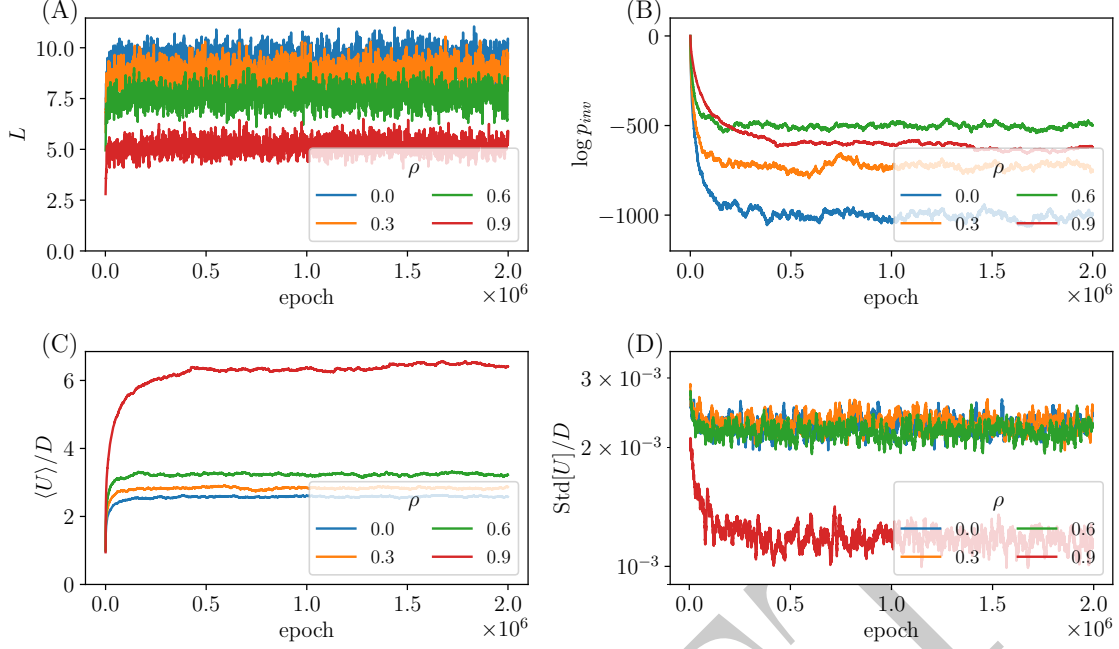
Figure 13. Effects of parent-mutant correlation, ρ, on evolution in CR model with $D = 50$, $\kappa = 0.4$. (A) Steady state $\overline{L}$ is a decreasing function of ρ. (B) $\overline{\log p_{inv}}$ is a nonmonotonic function of ρ, largest for intermediate ρ. (C) After initial transient, community-mean general fitness $\langle U \rangle$ saturates at steady state value. $\overline{\langle U \rangle}$ increases with ρ as distribution of extant U_i pushes farther into tail. (D) $\overline{\mathrm{Std}[U]}$ decreases with ρ.

we can Taylor expand the resource abundances, yielding

$$\mathcal{R}_\alpha^* \approx \frac{K_\alpha}{\mathcal{N}} \left(1 - \frac{\sum_j f_{j\alpha} n_j^*}{\mathcal{N}} \right), \qquad \text{(C1)}$$

where our definition of $\langle \mathcal{F} \rangle$ ensures that $\frac{1}{D} \sum_\alpha \mathcal{R}_\alpha^* = K_{\mathrm{tot}}/\mathcal{N} D$. If we approximate this expansion as an equality, at the fixed point the condition on the extant strains that their growth rates vanish becomes

$$0 = \sum_\beta \mathcal{G}_{i\beta} \frac{K_\beta}{\mathcal{N}} \left(1 - \frac{\sum_j f_{j\beta} n_j^*}{\mathcal{N}} \right) - 1 \text{ for } i \in \mathcal{C}$$

$$\text{(C2a)}$$

$$\implies 0 = \frac{1}{\mathcal{N}} \sum_\beta g_{i\beta} K_\beta - \frac{1}{\mathcal{N}^2} \sum_{\beta,j} g_{i\beta} K_\beta f_{j\beta} n_j^*$$

$$- \frac{\langle \mathcal{G} \rangle}{\mathcal{N}^2} \sum_{\beta,j} K_\beta f_{j\beta} n_j^* + \frac{\langle \mathcal{G} \rangle}{\mathcal{N}} \sum_\beta K_\beta - 1 \text{ for } i \in \mathcal{C},$$

$$\text{(C2b)}$$

where we have similarly defined $\langle \mathcal{G} \rangle \equiv \frac{1}{NK_{\mathrm{tot}}} \sum_{i,\beta} \mathcal{G}_{i\beta} n_i^* K_\beta$ and $g_{i\beta} = \mathcal{G}_{i\beta} - \langle \mathcal{G} \rangle$ so that the average of $\sum_\beta g_{i\beta} K_\beta$ over i (weighted by n_i^*) vanishes. This form allows us to separate out the general fitness and interaction pieces in the growth rate of each strain. On the right hand side of Equation C2b, the first term contains the general fitnesses and is proportional to $U_i - \langle U \rangle$. The second term is the contribution to the growth rate of strain i from the ecological interactions, and the last three terms are i-independent and constitute an effective Lagrange multiplier that keeps N roughly constant when $\langle \mathcal{G} \rangle \gtrsim \mathcal{O}(1)$.

In the assembled community, $\mathcal{N} = \mathcal{O}(L)$ and the $f_{j\alpha}$ and $g_{i\beta}$ have width σ_f and σ_g respectively. Therefore the

width of the general fitness term scales as $\sigma_g \sqrt{D}/L$ and the width of the interaction piece scales as $\sigma_f \sigma_g \sqrt{LD}/L^2$. Since $L = \mathcal{O}(D)$ for sufficiently large S, with each of the n_j^* being $\mathcal{O}(1)$, the distribution of general fitnesses in the assembled community with large S is a factor of $\sqrt{D}/\sigma_f$ wider than than the distribution of drives. Since σ_f and σ_g are $\lesssim 1$ in order to make the majority of entries in $\mathcal{F}$ and $\mathcal{G}$ positive, the general fitnesses dominate in the assembled community. When the community evolves, the $\mathcal{G}_i$ vectors elongate in the direction of the supply vector $\mathbf{K}$, which, as they push into their gaussian tail, decrease the width of the distribution of extant $\sum_\beta g_{i\beta} K_\beta / \mathcal{N}$, which is governed by the variations of the $\mathbf{g}_i$ vectors in the direction parallel to $\mathbf{K}$. However the width of the extant drive distribution does not decrease—this allows the distributions of general fitnesses and growth rates from ecological interactions to attain comparable widths in an evolved community (Section II B).

Appendix D: Algorithm to find fixed point

Integrating the ecological dynamics to find their stable fixed point is too slow to observe many of the features of the evolution that we discuss. In order to probe longer evolutionary trajectories, we have used a heuristic method to find the fixed point; we then check that it is stable and cannot be invaded by any of the extinct strains. Below we describe the algorithm for the linearized and LV models, assuming we are given some interaction matrix V, which can be made to include selective differences by adding s_i to the ith row of V (since $\sum_i \nu_i = 1$). We also comment
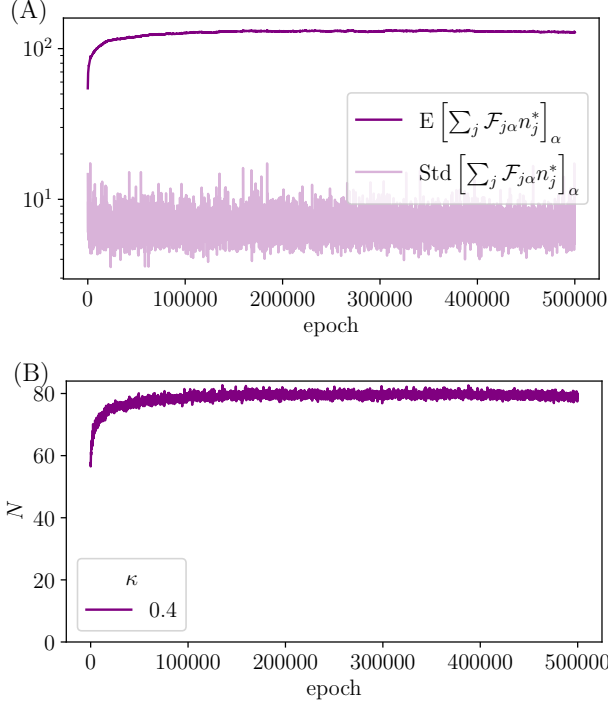
Figure 14. In Red Queen phase of the CR model ($D = 50$, $\kappa = 0.4$), linearization of resource abundances is justified, and total consumer population is roughly constant. (A) At steady state, consumption rates of resources, $\sum_j \mathcal{F}_{j\alpha} n_j^*$, have a narrow distribution, justifying expanding variations in this quantity out of the denominator. $K_\alpha = 1 \; \forall \; \alpha$, so weighting means and standard deviations by K_α does not affect results. (B) The total consumer population at the fixed point, $N = \sum_j n_j^*$, is roughly constant in the Red Queen phase, motivating switching to fractional abundances.

below on a slight modification of the algorithm for the CR model. The main challenge is to find the set $\mathcal{C}$ of strains in the community that has nonzero abundance at the stable uninvadable fixed point. We accomplish this task with an iterative method that converges most of the time.

1. Initialize $\mathcal{C}$ as the set of all strains in the community (including the new invader if applicable) and initialize the extinct set $\mathcal{E}$ as an empty set. V^* is the submatrix of interactions for the extant strains, and therefore contains the rows and columns corresponding to strains in $\mathcal{C}$.
2. Invert V^* to find a fixed point of the dynamics via $\nu_i^* = \Upsilon^* \sum_j V^{*-1}_{ij}$ and $\sum_i \nu_i^* = 1$ which fixes Υ^*. In general some of the ν_i^* will be negative. Remove the set $\{i | \nu_i^* < 0\}$ from $\mathcal{C}$ and add it to $\mathcal{E}$.
3. Calculate the invasion fitnesses ξ_j of each strain in $\mathcal{E}$ as $\xi_j = \sum_{k \in \mathcal{C}} V_{jk}\nu_k^* - \Upsilon^*$. Transfer the set of strains $\{j | \xi_j > 0\}$ from $\mathcal{E}$ to $\mathcal{C}$.
4. Iterate steps 2 and 3 until convergence to sets $\mathcal{C}$ and $\mathcal{E}$ such that $\nu_i^* > 0$ for all $i \in \mathcal{C}$ and $\xi_j < 0$ for all $j \in \mathcal{E}$.

One can then check stability of the final fixed point by computing the Jacobian of the dynamical equation. The vast majority of the time, the fixed point found in this way is stable.
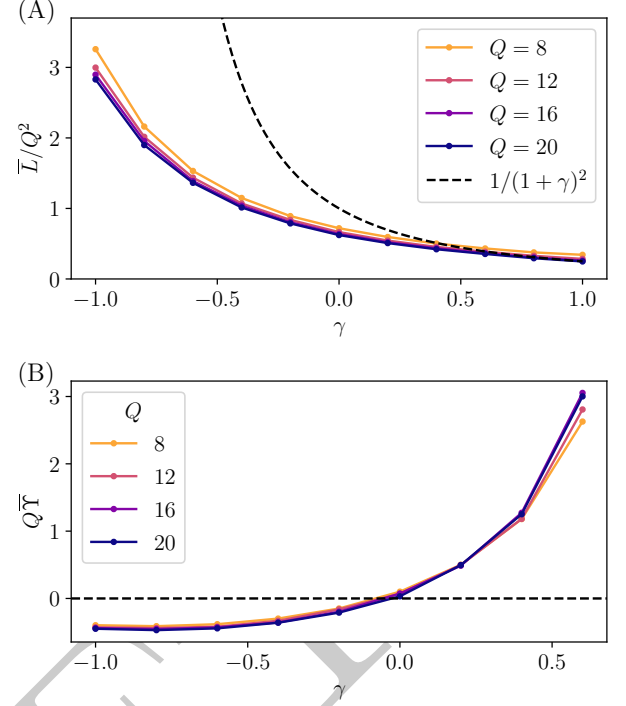

Figure 15. Evolutionary dynamics of LV model (with $\Sigma = 0$) depend on γ, with Q setting an overall scale for both $\overline{L}$ and $\overline{\Upsilon}$. (A) Steady state $\overline{L}$ in the LV model for $\Sigma = 0$ as a function of γ, scaled by Q^2. Dashed line shows $\hat{L}/Q^2 = 1/(1 + \gamma)^2$ from saturated assembled community [20]. When $\gamma > \gamma_c \approx 0.65$ the Red Queen phase breaks down, though L remains roughly constant, and its average is plotted. (B) $Q\overline{\Upsilon} = \mathcal{O}(1)$ across a range of $\gamma < \gamma_c$ in the Red Queen phase. For $\gamma > \gamma_c$ there is no steady state, and $\overline{\Upsilon}$ is formally infinite. Dashed line shows $\hat{\Upsilon} = 0$ from saturated assembled community.

When the algorithm does not converge within 10 iterations, we stop adding extinct strains with positive invasion fitnesses back into the community, and thereby converge to a fixed point with positive abundances, which is *invadable* by some of the strains that were declared extinct. This adjustment is most often deployed when $\mathcal{E}$ and $\mathcal{C}$ have similar sizes at the fixed point, which can happen for an initial assembled community with large S, or for the evolved state with small κ where L is small and each invader can cause $\mathcal{O}(L)$ extinctions. In these cases the iterative algorithm does not converge reliably. However for small perturbations to the community (as in the case of single invaders with large L), $\mathcal{E}$ is much smaller than $\mathcal{C}$ and the algorithm converges rapidly. Results from the error-prone regimes of evolution do not play a large part in any of our conclusions.

In the CR model, one cannot simply invert a matrix to find the fixed point for a subset of strains due to the nonlinearity in the interaction. Therefore we use a numerical solver to find the fixed point for a given $\mathcal{C}$—the rest of the algorithm is the same as in the linearized and LV cases.

Appendix E: LV evolution across range of γ

As in the linearized model, the LV model (shown in Figure 15 for $\Sigma = 0$) exhibits a Red Queen steady state in which $\overline{L}$ is less than its maximal value in an assembled community, $\hat{L} = Q^2/(1+\gamma)^2$, independent of Σ, as reviewed in Appendix F. The gap between $\hat{L}$ and $\overline{L}$ diverges as $\gamma \to -1$, reflecting the fact that the fragility Ξ is numerically larger for $\gamma \to -1$, and so it drives an average of one strain extinct per invasion for L farther from $\hat{L}$. For $\Sigma = 0$ (as in these simulations) $\overline{\Upsilon}$ changes sign at $\gamma = 0$, at which we have a quantitative understanding of the Red Queen phase (Section IV C). As in the linearized model, we expect weak dependence of $\overline{L}$ on Σ, and $\overline{\Upsilon} \sim \Sigma^2$ for gaussian $p(s)$ with large Σ.

In the oligarch phase, for $\gamma > \gamma_c$ with $\gamma_c \approx 0.65$ for $\Sigma = 0$, there are strains which grow to comprise an $\mathcal{O}(1)$ fraction of the population, and Υ continually increases with the community becoming harder and harder to invade. The width of the crossover between these two phases appears to sharpen for increasing Q (not shown), indicating a sharp transition in the high-diversity limit $Q \to \infty$. The oligarch phase is not a steady state and Υ runs away to arbitrarily large values, but L remains roughly constant, and it is this L that is plotted as $\overline{L}$ for $\gamma > \gamma_c$ in Figure 15A.

The rate of increase of Υ jumps discontinuously up from zero through the transition. This discontinuity suggests a first order transition, motivating our investigation of hysteresis around the critical γ. One can start the evolution with $\gamma > \gamma_c$, and, after some period of time, reduce γ to below γ_c for the ensemble from which newly invading strains are drawn. Figure 16A shows strain abundance dynamics during such a "quench." Here $\gamma = 1$ for the first 3×10^4 epochs of evolution, after which the invading strains start to enter the community with $\gamma = 0.5 < \gamma_c$. The strains into which an $\mathcal{O}(1)$ fraction of the community have condensed start to decrease in abundance, but instead of going extinct and turning over, they reach an apparent steady state in which they fluctuate around their still-$\mathcal{O}(1)$ mean values. This behavior becomes rarer as the post-quench γ is decreased, indicating a region of phase coexistence around γ_c. Figures 16B–D show the dynamics of L, Υ and p_{inv} through this process, compared with the behaviors for $\gamma = 0.5$ and $\gamma = 1$ constant throughout the evolution. If one looks only at L, the hysteresis is not apparent, with the post-quench L similar to L for a community that has always had $\gamma = 0.5$. Is it possible that, as for $\gamma = 0$, the fragility only depends on L but not the abundance distribution? If so, the fragility would set L in steady state.

In contrast to L, Υ and p_{inv} show clearly that the post-quench steady state is significantly different from a Red Queen state that has been evolving with $\gamma = 0.5$ from its beginning. This post-quench steady state displays properties of both the oligarch phase and the Red Queen phase, and may therefore be a useful to better understand the oligarch phase.

Appendix F: Cavity calculation for assembled community

Here we carry out the static cavity calculation for the assembled community in the linearized model: our results and approach are similar to that of ref. [21]. An analogous calculation for the LV model proceeds in a similar manner [19, 20, 23], and we summarize it at the end of this section. We consider a large community of strains at its fixed point, and add in another strain and resource, treating these as a small perturbation and then enforcing self-consistency between the statistics of this newly added strain and resource and the rest of the community. We will use the convention that repeated indices are summed over, unless otherwise stated.

Linearized model: The strains obey the dynamical equation

$$\frac{\dot{\nu}_i}{\nu_i} = s_i + V_{ij}\nu_j - \Upsilon \qquad \text{(F1)}$$

for $j \in \{1,\ldots,S\}$, where $V_{ij} = -G_{i\alpha}F_{j\alpha}$ for $\alpha \in \{1,\ldots,D\}$ and $\mathrm{E}[G_{i\alpha}F_{j\beta}] = \kappa\delta_{ij}\delta_{\alpha\beta}$. Υ is defined as $\Upsilon = \nu_i(s_i + V_{ij}\nu_j)$. Introducing resources $R_\alpha = -F_{j\alpha}\nu_j$, the fixed point abundances satisfy

$$0 = s_i + G_{i\alpha}\tilde{R}_\alpha - \Upsilon, \quad \tilde{R}_\alpha = -F_{j\alpha}\tilde{\nu}_j, \qquad \text{(F2)}$$

where tildes indicate the fixed point values prior to adding a new strain and resource. Now we add in a new strain and resource with fixed point abundances ν_0 and R_0 respectively. Defining quantities $\zeta_0 = G_{0\alpha}\tilde{R}_\alpha$ and $\phi_0 = -F_{j0}\tilde{\nu}_j$, we have

$$0 = s_0 + \zeta_0 + G_{0\alpha}\delta R_\alpha - \Upsilon, \quad R_0 = \phi_0 - F_{j0}\delta\nu_j. \quad \text{(F3)}$$

If we define local susceptibilities $\eta_\alpha = \frac{dR_\alpha}{d\phi_\alpha}$ and $\chi_i = \frac{d\nu_i}{d\zeta_i}$, then the changes in R_α and ν_j are $\delta R_\alpha = -\eta_\alpha F_{0\alpha}\nu_0$ and $\delta\nu_j = \chi_j G_{j0}R_0$. (Here repeated indices are not summed.)

Assuming self-averaging (distributions concentrate around their mean for large systems), the new strain and resource abundances are

$$\nu_0 = \frac{\zeta_0 + s_0 - \Upsilon}{\kappa H}, \quad R_0 = \frac{\phi_0}{1+\kappa X}, \qquad \text{(F4)}$$

where $H = \sum_\alpha \eta_\alpha$ and $X = \sum_i \chi_i$ are global susceptibilities.

Since ζ and ϕ are weighted sums of the G and F matrix elements respectively, they are gaussian distributed with mean 0. We will denote their variances as σ_ζ^2 and σ_ϕ^2 respectively. Let us take the s to be gaussian distributed with mean 0 and variance Σ^2, and define $\sigma_\xi^2 = \sigma_\zeta^2 + \Sigma^2$ and $Y = \Upsilon/\sigma_\xi$. We can now write self-consistency conditions for the variables σ_ϕ, σ_ξ, H, X and Y, which require that the distributions of the added strain and resource abundances are the same as the distributions in the original community: in other words, Equation F4 is true for all strains, not just the cavity strain. Defining partial moments

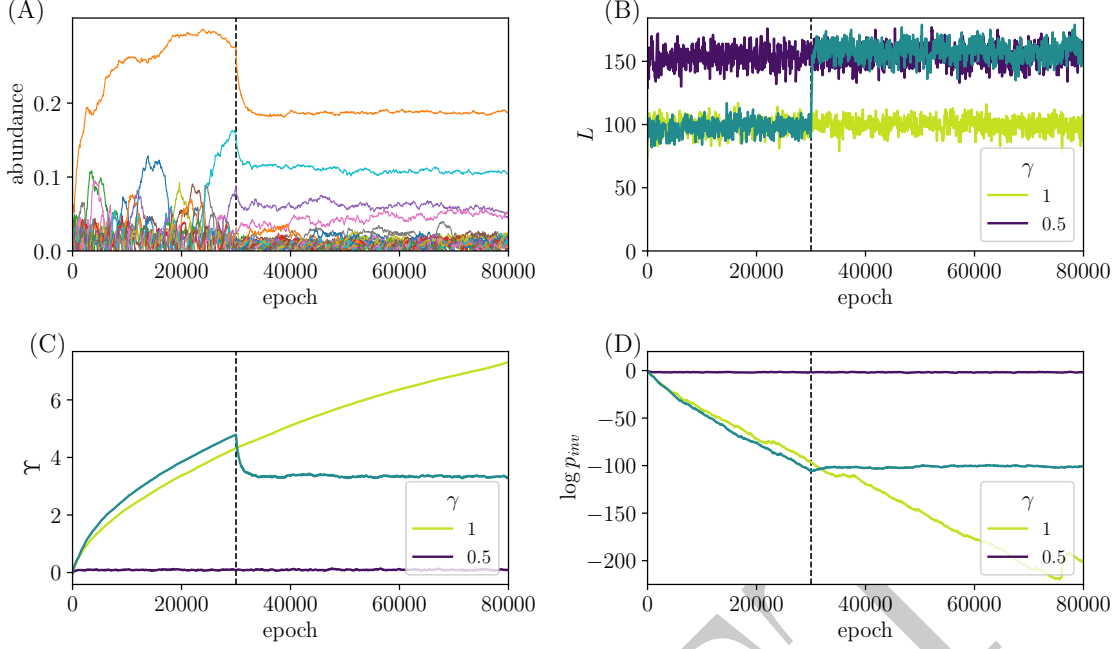$$W_m(Y) = \frac{1}{\sqrt{2\pi}}\int_0^\infty x^m e^{-(x+Y)^2/2}dx, \qquad \text{(F5)}$$

Figure 16. Transition between Red Queen and oligarch phases is first order, and exhibits hysteresis and phase coexistence when the system is tuned across the transition, for $Q = 20$, $\Sigma = 0$. (A) Strain trajectories where initially $\gamma = 1$, and after 3×10^4 epochs (at the vertical dashed line) γ is "quenched" to $0.5 < \gamma_c$. System does not revert back to the original steady state, instead showing several long-lived strains persisting. (B) The trajectory of L over such an evolution (dark green), shown alongside L for constant $\gamma = 0.5$ and $\gamma = 1$. The value of L is indistinguishable in the post-quench phase (after a short transient) from the $\gamma = 0.5$ Red Queen phase. (C)–(D) Dynamics of Υ and p_{inv} show a strong departure after the quench from their values when $\gamma = 0.5$ throughout.

our self-consistency conditions are

$$X = \frac{S}{\kappa H} W_0(Y) \qquad \text{(F6a)}$$

$$1 = \frac{S\sigma_\xi}{\kappa H} W_1(Y) \qquad \text{(F6b)}$$

$$\sigma_\phi^2 = \frac{S\sigma_\xi^2}{\kappa^2 H^2} W_2(Y) \qquad \text{(F6c)}$$

$$H = \frac{D}{1 + \kappa X} \qquad \text{(F6d)}$$

$$\sigma_\xi^2 = \frac{D\sigma_\phi^2}{(1 + \kappa X)^2} + \Sigma^2. \qquad \text{(F6e)}$$

These equations can be reduced to

$$X = \frac{SW_0(Y)}{\kappa(D - SW_0(Y))} \qquad \text{(F7a)}$$

$$H = D - SW_0(Y) \qquad \text{(F7b)}$$

$$\sigma_\xi = \kappa \frac{D - SW_0(Y)}{SW_1(Y)}. \qquad \text{(F7c)}$$

along with an equation for Y alone:

$$\frac{SW_2(Y)}{\kappa^2 D} = 1 - \left[\frac{\Sigma SW_1(Y)}{\kappa(D - SW_0(Y))} \right]^2. \qquad \text{(F8)}$$

A solution to these equations give us the distributions of the strain and resource abundances in the community, as well as the susceptibilities H and X, and the scaled Lagrange multiplier Y. This mean field solution is expected to describe the unique stable uninvadable fixed point as

long as it is stable to a small random perturbation on each of the strains' growth rates.

To look at the nonlinear instability that signals the breakdown of the mean field theory, we add a random field h_i to each of the biases and look at the change $\delta\xi_0$ in the bias of strain 0. We find that

$$\delta\xi_0 = \eta_\alpha G_{0\alpha} F_{j\alpha} \chi_j \delta\xi_j + h_0, \qquad \text{(F9)}$$

where the sum is over both α and j. Then in evaluating the mean square of both sides of this equation we have

$$\mathrm{E}\left[(\eta_\alpha G_{0\alpha} F_{j\alpha} \chi_j \delta\xi_j)^2\right] = \sum_\alpha \eta_\alpha^2 \sum_j \chi_j^2 \, \mathrm{E}[\delta\xi_i^2]. \qquad \text{(F10)}$$

Assuming that the strains are statistically identical, we can drop the strain indices, to get

$$\mathrm{E}[\delta\xi^2] = \frac{\mathrm{E}[h^2]}{1 - \sum_\alpha \eta_\alpha^2 \sum_j \chi_j^2} \qquad \text{(F11a)}$$

$$\implies \sum_i \mathrm{E}[\delta\nu_i^2] = \frac{\sum_j \chi_j^2 \, \mathrm{E}[h^2]}{1 - \sum_\alpha \eta_\alpha^2 \sum_j \chi_j^2} = \Xi \times \mathrm{E}[h^2] \quad \text{(F11b)}$$

where we have defined the *fragility* Ξ as the factor relating the variance of the random field to the sum of the squared changes in abundances that it produces. Since $\eta_\alpha = H/D$ and $\chi_j = X/SW_0(Y)$ in the mean field solution, we can write the fragility in terms of $L = SW_0(Y)$ as

$$\Xi = \frac{L/\kappa^2 H^2}{1 - L/\kappa^2 D} = \frac{LD}{(D - L)^2(\kappa^2 D - L)}. \qquad \text{(F12)}$$

An instability signaling a phase transition occurs when the fragility diverges, at $L = \kappa^2 D$. We could have (more naturally) defined the fragility as the response of the strain abundances to adding a random field to the resource fields $\{\phi_\alpha\}$, but the form of divergence at the phase transition is the same in both cases (and in LV model, only the fragility we use here is defined.) The location of the divergence implies that the number of surviving strains is $\hat{L} = \kappa^2 D$ at the transition. Substituting this into Equation F8 tells us that $\hat{Y}$ is the solution to

$$\frac{W_2(\hat{Y})}{W_0(\hat{Y})} = 1 - \left[\frac{W_1(\hat{Y})\Sigma\kappa}{W_0(\hat{Y})(1-\kappa^2)}\right]^2. \qquad (F13)$$

For $\Sigma = 0$, the solution to this equation is $\hat{Y} = 0$ and $\hat{S} = 2\kappa^2 D$. For $\Sigma \gtrsim 1$, we have $\hat{Y} \approx \frac{\Sigma\kappa}{1-\kappa^2}$ and $\hat{\sigma}_\xi \approx \Sigma$, meaning that $\hat{\Upsilon} \approx \frac{\kappa\Sigma^2}{1-\kappa^2}$ and $\hat{S} \approx \sqrt{2\pi}\kappa^2 D\hat{Y}e^{\hat{Y}^2/2} \sim \exp[\frac{\kappa^2\Sigma^2}{2(1-\kappa^2)^2}]$. The reciprocal of $\hat{S}$ gives the invasion probability of the saturated community by an independent invader. The scaling of $\hat{\Upsilon}$ and $\hat{S}$ with Σ and κ is reflected in the scaling of $\overline{\Upsilon}$ and $\overline{\log p_{inv}}$ in the Red Queen steady state.

Note that the linear susceptibility X diverges when $L = D$. The nonlinear instability occurs for lower S than the linear instability for $\kappa < 1$, meaning that the cavity solution becomes invalid before it displays a linear instability. The nonlinear and linear instabilities occur in the same place for $\kappa = 1$, signaling the onset of the "shielded" phase when $\Sigma = 0$ [35]. When $\Sigma > 0$, $\hat{S}$ diverges for $\kappa \to 1$, indicating the marginality of the shielded phase.

LV model: We now summarize the cavity calculation for the LV model. In this case the cavity method yields, for the fixed point abundance of the cavity strain,

$$\nu_0 = \frac{\zeta_0 + s_0 - \Upsilon}{Q - \gamma X} \qquad (F14)$$

where $\zeta_0 = \sum_j V_{0j}\tilde{\nu}_j$, and $X = \sum_i \chi_i = \sum_i \frac{d\nu_i}{d\zeta_i}$. Note the similarity with the expression from the linearized model

$$\nu_0 = \frac{\zeta_0 + s_0 - \Upsilon}{\kappa D - \kappa^2 HX}, \qquad (F15)$$

(a consequence of Equations F4 and F7) if one identifies $\kappa\sqrt{D}$ with Q, κ^2 with γ and HX with $\sqrt{D}X$, which account for the integrated-out resources as well as the factor of $\sqrt{D}$ between the LV and linearized models.

The self consistency equations for the LV model with gaussian distributed s_i of variance Σ^2, for unknowns σ_ξ, Y and X, are then

$$X = \frac{S}{Q - \gamma X}W_0(Y) \qquad (F16a)$$

$$1 = \frac{S\sigma_\xi}{Q - \gamma X}W_1(Y) \qquad (F16b)$$

$$\sigma_\xi^2 = \frac{S\sigma_\xi^2}{(Q - \gamma X)^2}W_2(Y) + \Sigma^2 \qquad (F16c)$$

where, as before, $Y = \Upsilon/\sigma_\xi$. These equations can be re-

duced to

$$X = \frac{Q - \sqrt{Q^2 - 4\gamma SW_0(Y)}}{2\gamma} \qquad (F17)$$

$$\sigma_\xi = \frac{Q + \sqrt{Q^2 - 4\gamma SW_0(Y)}}{2SW_1(Y)} \qquad (F18)$$

along with an equation for Y alone:

$$\left[\frac{Q + \sqrt{Q^2 - 4\gamma SW_0(Y)}}{2}\right]^2 = SW_2(Y) + (\Sigma SW_1(Y))^2. \qquad (F19)$$

The fragility can be calculated from local susceptibility $\chi_i = 1/(Q - \gamma X)$ in terms of $L = SW_0(Y)$:

$$\Xi = \frac{L}{(Q - \gamma X)^2 - L} \qquad (F20)$$

implying a divergence of Ξ at $\hat{L} = \frac{Q^2}{(1+\gamma)^2}$.

Appendix G: Exponential distribution of general fitnesses

In order to analyze the transient behavior of L as $\Delta(\Upsilon)$ (the local scale of the s_i distribution) decreases during the evolutionary transient, it is useful to investigate the behavior of the assembled community with an exponential distribution of positive s_i given by $p(s) = \frac{1}{\Delta}e^{-s/\Delta}\Theta[s]$. We define the variable $x_i \equiv s_i + \zeta_i$. The self consistency conditions then deal with the distribution of x_i, which is

$$p_x(x) = \frac{1}{\Delta}\exp\left[-\frac{x}{\Delta} + \frac{\sigma_\zeta^2}{2\Delta^2}\right]\Phi\left(\frac{x}{\sigma_\zeta} - \frac{\sigma_\zeta}{\Delta}\right), \qquad (G1)$$

where Φ is the standard normal CDF. In particular, if we define partial moments

$$\tilde{W}_m(\Upsilon, \sigma_\zeta) = \int_0^\infty y^m p_x(y + \Upsilon)dy \qquad (G2)$$

then our self consistency conditions can be written down in a similar fashion to those for the gaussian $p(s)$ in Equation F6. For the LV model, these conditions are

$$X = \frac{S}{Q - \gamma X}\tilde{W}_0(\Upsilon, \sigma_\zeta) \qquad (G3)$$

$$1 = \frac{S}{Q - \gamma X}\tilde{W}_1(\Upsilon, \sigma_\zeta) \qquad (G4)$$

$$\sigma_\zeta^2 = \frac{S}{(Q - \gamma X)^2}\tilde{W}_2(\Upsilon, \sigma_\zeta) \qquad (G5)$$

For large Υ, the lower limit of the exponential $p(s)$ does not matter and the distribution $p_x(x)$ is roughly exponential. Therefore we have

$$\tilde{W}_m(\Upsilon, \sigma_\zeta) \approx m! \times \Delta^m \exp\left[-\frac{\Upsilon}{\Delta} + \frac{\sigma_\zeta^2}{2\Delta^2}\right], \qquad (G6)$$

which allows us to write the self consistency conditions just in terms of the number of extant strains $L = \tilde{W}_0(\Upsilon, \sigma_\zeta)$,
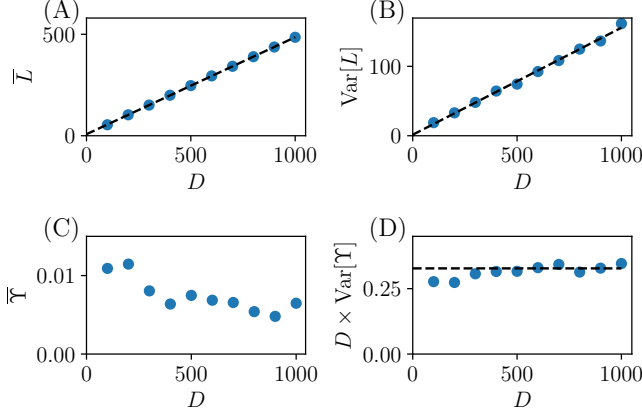
Figure 17. Fluctuations of L and Υ in the Red Queen phase of the linearized model with $\kappa = 0.8$, $\rho = 0$ and $\Sigma = 0$. (A)–(B) Both the mean and variance of L scale roughly linearly with D, so that fluctuations are smaller than the mean by a factor of $\sqrt{D}$. (C)–(D) The mean Υ in steady state is small with ambiguous dependence on D, with the variance of Υ scaling as $1/D$.

regardless of S and the normalization of $p(s)$:

$$X = \frac{L}{Q - \gamma X} \tag{G7}$$

$$1 = \frac{\Delta L}{Q - \gamma X} \tag{G8}$$

$$\sigma_\zeta^2 = \frac{2\Delta^2 L}{(Q - \gamma X)^2} \tag{G9}$$

In the large-Δ regime, we therefore have $X \approx 1/\Delta$, implying that to leading order in Δ, the diversity is given by $L(\Delta) \approx Q/\Delta$. A similar analysis in the linearized model yields $L(\Delta) \approx \kappa D/\Delta$. This analysis describes the behavior for $p(s)$ decaying faster than exponentially as well, where Δ is a function of Υ and the distribution of s_i has pushed far enough into the tail of $p(s)$ to look locally exponential around Υ.

Note that $\hat{\Upsilon}$ and $\hat{S}$ diverge for sufficiently large Δ, which scales as $(1 + \gamma)/Q$ in the LV model and $1/\kappa$ in the linearized model. This is because as Δ increases, the maximum L—given by $X(Q - \gamma X)$ and $X\kappa D/(1 + \kappa X)$ in the LV and linearized models respectively—becomes smaller than

$Q^2/(1 + \gamma)^2$ or $\kappa^2 D$ respectively, and so Ξ cannot diverge. Therefore sufficiently large Δ eliminates the unstable phase for an exponential $p(s)$.

Appendix H: Fluctuations in steady state

Figures 17A and B show that in the Red Queen steady state for the linearized model, L fluctuates around a mean value $\overline{L}$ which is proportional to D, with fluctuations that scale with $\sqrt{D}$. This scaling is consistent with expectations from the finite fragility of the Red Queen steady state—with evolutionary perturbations to the community not causing large changes and fluctuations therefore down by a factor of $1/\sqrt{D}$. Figures 17C and D show the same statistics for Υ. $\overline{\Upsilon}$ is naively $\mathcal{O}(1)$ but is anomalously small in numerics (since $\Sigma = 0$), and it is possible that the coefficient of its $\mathcal{O}(1)$ part is zero, with a $1/\sqrt{D}$ correction showing up in numerics. Fluctuations in Υ have size $\mathcal{O}(1/\sqrt{D})$. For $\Sigma = 0$, we know that $\overline{\Upsilon} = 0$ in the limit $\kappa \to 0$, $D \to \infty$, $\kappa^2 D = Q^2$ (which reduces to the $\gamma = 0$ Lotka-Volterra model), as well as when $\kappa \to 1$, since in the latter case Υ becomes nonpositive and asymptotes to 0 from below. It is therefore possible that $\overline{\Upsilon} = 0$ for intermediate κ as well, but our results are inconclusive.

Appendix I: Dynamical cavity analysis in simplest case

Here we carry out the dynamical cavity analysis for the evolutionary steady state of the LV model with $\gamma = 0$, $\Sigma = 0$. In particular we calculate the temporal correlation function of the cavity strain in Equation 10 and derive self-consistency conditions to find the statistical properties of the evolution of strain abundances. Throughout, we suppress the index 0 of the cavity strain.

Solution for correlation function: Because of the absorbing boundary condition at $\nu = 0$, we can use the solution of the Ornstein Uhlenbeck process (Equation 11) together with the method of images to write down the distribution of $\nu(T)$ given that it started at some initial value $\nu(\mathbb{T})$ at invasion time $T = \mathbb{T}$. For simplicity we set $\mathbb{T} = 0$: integrating the correlation function over $\mathbb{T}$—as we will later do—is the same as integrating over T. We use the fact that, for $\gamma = \Sigma = \overline{\Upsilon} = 0$, we have $\nu = \zeta/Q$. Therefore we can write the distribution of $\nu(T)$ conditioned on $\nu(0)$ as

$$p^*(\nu(T)|\nu(0)) =$$
$$\frac{1}{\sqrt{2\pi}\sigma(T)} \left[\exp\left(-Q^2 \frac{[\nu(T) - \nu(0)e^{-bT}]^2}{2\sigma(T)^2} \right) - \exp\left(-Q^2 \frac{[\nu(T) + \nu(0)e^{-bT}]^2}{2\sigma(T)^2} \right) \right], \text{ where } \sigma^2(T) = \frac{J}{b}(1 - e^{-2bT}), \tag{I1}$$

where we have used the notation p^* to denote the probability density with an absorbing boundary condition at $\nu = 0$: the integral of p^* over $\nu(T)$ is the survival probability of the strain as a function of time T. We calculate the expectation $\langle \nu(T)\nu(T + \tau) \rangle$ over the joint distribution

$p^*(\nu(T), \nu(T+\tau)|\nu(0)) = p^*(\nu(T)|\nu(0)) \times p^*(\nu(T+\tau)|\nu(T))$. We then integrate over all times $T \in [0, \infty)$ and over all

$\nu(0) \in [0, \infty)$ distributed according to

$$p(\nu(0)) = \frac{r}{\sqrt{2\pi}\tilde{\sigma}} \exp\left[-\frac{Q^2\nu(0)^2}{2\tilde{\sigma}^2}\right] \Theta[\nu(0)], \quad (I2)$$

where we have introduced $\tilde{\sigma} = \sqrt{J/b}$ as the standard deviation of the incoming drive distribution, and r as the rate of attempted invasions, which is 2 in our case since $\overline{\Upsilon} = 0$ so each invader has probability $1/2$ of successfully invading. The integrals over T and $\nu(0)$ correspond to summing over all the strains in the community and can be done exactly, yielding

$$C(\tau) = \mathrm{E}[\langle\nu(T)\nu(T+\tau)\rangle]_{T,\nu(0)} \quad (I3a)$$

$$= \frac{r\tilde{\sigma}^2}{4Q^2b}(1 + 2\log 2)e^{-b\tau}. \quad (I3b)$$

We now assert that the dynamics of the cavity drive are the same as those of the drives in the community. Since the cavity drive can fluctuate in sign, its correlation function is $C(\tau) = \tilde{\sigma}^2 e^{-b\tau}$, from an unconditioned Ornstein Uhlenbeck process at stationarity. Equating our two expressions for $C(\tau)$ gives

$$1 = \frac{r}{4Q^2b}(1 + 2\log 2). \quad (I4)$$

We can get another self-consistency condition by enforcing $\sum_j \nu_j = 1$. Note that $\langle\nu(T)|\nu(0)\rangle = \nu(0)e^{-bT}$, where the expectation is taken with an absorbing boundary condition at $\nu = 0$. Then integrating this over the past and over the half-gaussian distribution of starting positions (as before) gives $r\tilde{\sigma} = Qb\sqrt{2\pi}$.

$\overline{L}$, the number of strains in steady state, is determined by the fact that the survival probability up to time T is $\langle\Theta[\nu(T)]|\nu(0)\rangle = \mathrm{erf}[Q\nu(0)e^{-bT}/\sqrt{2}\sigma(T)]$. Integrating over all past times and over the distribution of $\nu(0)$ gives $\overline{L} = \frac{r}{2b}\log 2$. Now we have two self-consistency conditions for b and $\tilde{\sigma}$, along with an equation for $\overline{L}$:

$$r\tilde{\sigma} = Qb\sqrt{2\pi} \quad (I5)$$

$$4Q^2\frac{b}{r} = 1 + 2\log 2 \quad (I6)$$

$$\overline{L} = \frac{r}{2b}\log 2. \quad (I7)$$

We can solve these equations in terms of Q. The most easily testable prediction from numerics is

$$\overline{L} = \frac{2\log 2}{1 + 2\log 2}Q^2 \approx 0.58 \times Q^2, \quad (I8)$$

which agrees well with numerics for $\overline{L}$ with $\gamma = 0$ across a range of Q (Figure 10A). We also find that $\tilde{\sigma} = \sqrt{2\pi}\lambda/Q$, and $b = \lambda r/Q^2$ for $\lambda = \frac{1 + 2\log 2}{4}$, implying that $J = 2\pi r\lambda^3/Q^4$. Therefore the theory prediction for the correlation function of the cavity drive is

$$C(\tau) = \frac{2\pi\lambda^2}{Q^2}\exp\left[-\frac{\lambda}{Q^2}r\tau\right]. \quad (I9)$$

As stated previously, $r = 2$ since this makes the integral of the half-gaussian incoming drive distribution (and therefore the rate of incoming strains) equal to 1.

Lifetime distribution: The distribution of strain lifetimes can be computed using known results from the theory of Ornstein Uhlenbeck (OU) processes—which can also be derived from the method of images. A useful result [59] is that the distribution of survival times $\mathcal{T}$ for an OU process $\dot{x} = -bx + \sqrt{2J}\eta(T)$ starting at position $x_0 > 0$ and ending when it hits the origin is

$$p_{\mathcal{T}}(\mathcal{T}|x_0) = \frac{2b^{3/2}x_0}{\sqrt{2\pi J}}\frac{e^{2b\mathcal{T}}}{(e^{2b\mathcal{T}} - 1)^{3/2}}\exp\left[-\frac{bx_0^2}{2J(e^{2b\mathcal{T}} - 1)}\right]. \quad (I10)$$

We can integrate over the distribution

$$p_{x_0}(x_0) = \frac{2\Theta(x_0)}{\sqrt{2\pi}\tilde{\sigma}}e^{-x_0^2/2\tilde{\sigma}^2}, \text{ where } \tilde{\sigma} = \sqrt{J/b}. \quad (I11)$$

We then obtain $p_{\mathcal{T}}(\mathcal{T}) = 2b/\pi\sqrt{e^{2b\mathcal{T}} - 1}$, as quoted in Equation 14.

Abundance distribution: The steady state distribution of abundances can also be computed from our self consistent solution. The abundance distribution obeys a Fokker Planck equation (with the appropriate boundary conditions) that captures the dynamics of the drives from Equation 11 and includes a *source* term for incoming invaders. Therefore in the steady state, the condition for stationarity of the distribution $P(\nu)$ is

$$\frac{b\tilde{\sigma}^2}{Q^2}P''(\nu) + b[P(\nu) + \nu P'(\nu)] + \frac{rQ}{\sqrt{2\pi}\tilde{\sigma}}\exp\left(-\frac{\nu^2Q^2}{2\tilde{\sigma}^2}\right) = 0 \quad (I12)$$

with boundary conditions $P(0) = P(\nu \to \infty) = 0$. Equation I12 contains effects of the systematic and random parts in $\delta\nu$ as well as the input from invading strains.

If we change variables to $z = Q^2\nu$ and $\Pi(z) = \frac{1}{Q^4}P(\nu)$ so that $\overline{L} = Q^2\int dz\Pi(z)$, we find that the number density $\Pi(z)$ of the rescaled abundances z obeys the Q-independent differential equation

$$2\pi\lambda^3\Pi''(z) + \lambda[\Pi(z) + z\Pi'(z)] + \frac{1}{2\pi\lambda}\exp\left(-\frac{z^2}{4\pi\lambda^2}\right) = 0 \quad (I13)$$

with $\lambda = \frac{1 + 2\log 2}{4}$. The solution to this equation with boundary conditions $\Pi(0) = 0$, $\Pi(\infty) = 0$ is plotted in Figure 9 as the theoretical prediction for the distribution of scaled abundances in the evolved community.

Appendix J: Evolutionary transients with $\Sigma > 0$

In Section IV E we discuss the transient behavior of the evolution as the distribution of extant s_i pushes into the tail of $p(s)$. For a gaussian distribution of the s_i, we expect that, during the transient,

$$\frac{\Upsilon}{\Sigma^2} \sim \left(\frac{T}{\Sigma^2}\right)^{1/3}. \quad (J1)$$

In Figure 18 we find that this prediction agrees well with numerics for Υ as a function of evolutionary time, and the collapse of the curves scaled by Σ verifies the predictions. The data indicate that, in agreement with observation in
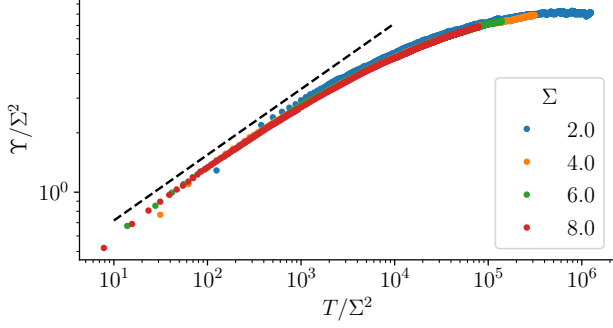
Figure 18. The transient behavior of Υ during the evolutionary dynamics with large Σ in the linearized model. Parameters are $D = 200$, $\kappa = 0.8$ and $\rho = 0$. Dashed black line has a slope of $1/3$. Each simulation was run for 5×10^6 epochs.

Figure 5C, for our choice of parameters, the simulations with $\Sigma \geq 4$ have not yet reached their steady state by the end of 5×10^6 epochs. The dynamics of L during the transient (not shown) are less well-fit by a simple power law since L varies over a smaller range and has a larger constant correction.

Appendix K: $C(\tau)$ for nonzero γ and Σ

We expect that our solution for the evolutionary dynamics of the LV model at $\gamma = 0$ remains qualitatively valid across a range of parameters. Here we check that assumption by examining the correlation function $C(\tau) = \sum_j \nu_j(T)\nu_j(T + \tau)$ for various Σ in the linearized model with $\kappa = 0.8$, and various γ in the LV model.

In Figure 19B we show that larger γ increases the timescale of turnover as evidenced by slower decay of $C(\tau)$—with time rescaled appropriately by $\overline{L}$, which is equal to the average strain lifetime. Increasing γ decreases turnover rate since invading strains will push existing strains to higher abundance and slow down their rate of extinction. However the correlation functions still look roughly exponential at large τ. For $\gamma < 0$ it appears that $C(\tau)$ decays slightly faster for small τ than for large τ, with the opposite trend for $\gamma > 0$.

In Figure 19A we see that increasing Σ in the linearized model has the effect of decreasing the rate of decay of $C(\tau)$. Here we expect that $C(\tau)$ decays as a Σ-dependent power law—for reasons we explain below—due to the presence of rare high-fitness strains and their long residence times in the community. In the inset of Figure 19A we see that the mean lifetime of a strain grows faster than exponentially with its general fitness, consistent with inverse-gaussian dependence, and indicating the large effect of strains with large s_i.

We can understand the behavior for $\Sigma > 0$ through a toy model for evolution by independent invaders with $\Sigma > 0$. In this toy model, the width of the extant s_i distribution is $\mathcal{O}(1)$ at steady state, as in the linearized model, in contrast to the $\mathcal{O}(1/Q)$ steady state width in the LV model (see Section III B).

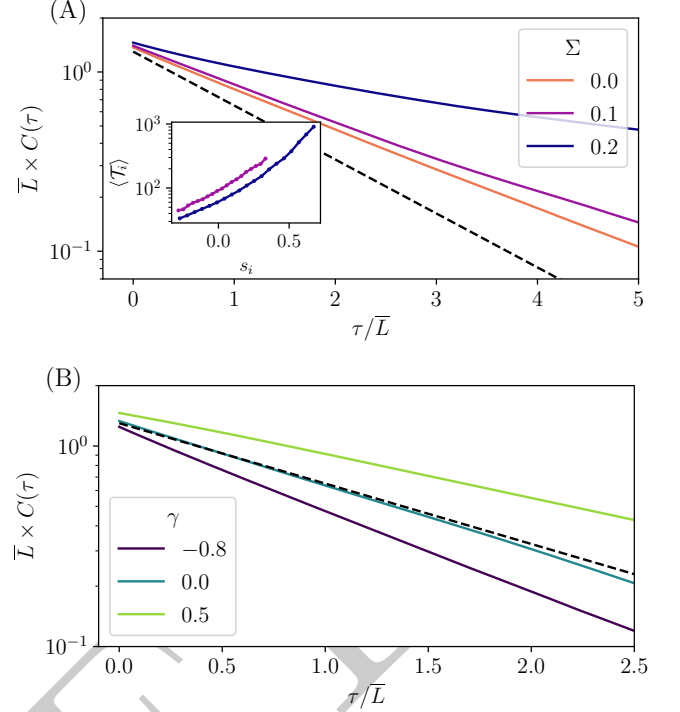We can try the same *Ansatz* that worked in the $\Sigma = 0$

Figure 19. The abundance correlation function $C(\tau)$ for various parameter choices. (A) $C(\tau)$ in the linearized model with $\kappa = 0.8$ decays slower with increasing Σ. Long transients make predicted asymptotic power law behavior difficult to observe. Inset shows that mean lifetime for a strain with general fitness s_i grows faster than exponentially—consistent with predicted gaussian-fast in s_i. Dashed black line shows correlation function of the $\gamma = 0$ LV model for comparison. (B) In the LV model with $\Sigma = 0$, $C(\tau)$ decays slower with increasing γ. Dashed black line is the simple-exponential prediction for $\gamma = 0$, with deviations from this result due to finite sampling in the numerics. In both panels τ is normalized by basic timescale of the community turnover $\overline{L}$.

and $\gamma = 0$ case, namely that the drives obey an Ornstein Uhlenbeck process $\dot{\zeta} = -u\zeta + \sqrt{2J}\eta(T)$, but now the absorbing boundary condition for strain i is at $\zeta = \overline{\Upsilon} - s_i$. Therefore strains with a sufficiently above-average s_i only go extinct when their drive becomes negative, which takes a time gaussian-long in $\overline{\Upsilon} - s_i$. In particular, defining $y_i = \overline{\Upsilon} - s_i$, we expect that the characteristic residence time in the community for strain i is $\langle \mathcal{T}_i \rangle \sim \frac{y_i^2}{J} e^{u y_i^2/4J}$. At large τ, the decay of the correlation function $C(\tau)$ will be roughly the probability that a strain survives for time τ—therefore, as stated in Section IV E, we need to integrate over the distribution of the y_i in order to get

$$C(\tau) \sim \int dy\, p(y) \exp\left[-\frac{J\tau}{y^2 e^{u y^2/4J}} \right]. \qquad (K1)$$

For $p(y) \sim e^{-\frac{1}{\psi}(y/\Sigma)^\psi}$, we find that

$$C(\tau) \sim \exp\left[-\frac{1}{\psi}\left(\frac{2\sqrt{J/u}}{\Sigma} \right)^\psi \log^{\psi/2}\left(\frac{\Sigma^\psi u\tau}{2} \right) \right]. \qquad (K2)$$

Therefore we see that an exponential *Ansatz* for correlation function of the dynamics is not self-consistent. In particu-

lar, this *Ansatz* yields a $C(\tau)$ that decays as a Σ-dependent power law for a gaussian distribution of s (corresponding to $\psi = 2$). This result will change prefactors in the scaling of $\log\langle\mathcal{T}_i\rangle$ with y_i, but this scaling will still be quadratic due to the gaussian steady state of the unconditioned drive dynamics—and corresponding truncated gaussian distribution for the invader drive.

We expect that this general mechanism for producing long tails in the lifetime distribution—and therefore $C(\tau)$—is at work whenever $\Sigma > 0$. Having an upper limit for the s distribution (and therefore the y distribution) will cut off the power law at τ larger than the maximum $\langle\mathcal{T}_i\rangle$, beyond which the correlation function will look exponential. We have not pursued quantitatively testing these forms for $C(\tau)$ against our numerics due to unknown numerical factors that would affect the power scaling, as well as long transients that are expected to arise in $C(\tau)$. However, the results in Figure 19B do clearly indicate a broadening of $C(\tau)$ with increasing Σ.

In order for the saddle point approximation that yields Equation K2 to be accurate, τ should be gaussian-large in Σ. For small τ the behavior will be dominated by y_i's drawn from the small-s part of the s_i distribution and will therefore look exponentially distributed—therefore the behavior $\log C(\tau) \sim -\sqrt{\log \tau}$ for an exponential $p(s)$ might be better testable in simulations as there is no obvious transient for this behavior.

In the cases where the strain lifetime distribution has diverging mean, we expect that the equality between mean lifetime and mean diversity breaks down from effects of the memory kernel, as described in Section IV E. However, more work is needed to understand this mechanism in detail.

Appendix L: Comparison with results of de Pirey and Bunin

Recent work of de Pirey and Bunin [51] (dPB for short) has explored a Red Queen-like phase in a different context. dPB study *ecological dynamics* in a $\gamma = 0$ Lotka-Volterra model in the limit of small migration for each species, such that there is a clear separation between species which are extant and species which would be extinct if not for the small migration. They find that if the diversity is high enough, the ecological dynamics are chaotic such that there are species which go nearly extinct before rising again to high abundance and rejoining the extant subset which was almost at a stable fixed point. Due to a separation of timescales arising from the very small migration rate, the ecological dynamics therefore consist of a series of fixed points punctuated by changes that occur when a near-extinct species rises up to non-negligible abundance. When a species transitions from near-extinct to extant, the process is similar to an independent invader joining the community, as the abundance of the invader can jump up to a nonzero abundance determined by its ecological interactions with the extant community. This entrance to the community causes a change in the biases of the other strains in the community which can drive low abundance strains out and therefore produces a steady state in which the abundance distribution is depleted at low abundance as compared to the truncated gaussian of the assembled community—indeed dPB find that their distribution at steady state is not well approximated by a truncated gaussian, and the number of extant species at any given time is less than the total capacity of an assembled community.

However, in the model studied by dPB, the steady state abundance distribution does not vanish at zero abundance. This is because when a species goes from extant to nearly-extinct, its probability of rejoining the extant community due to an ecological fluctuation is a decreasing function of time as ecological fluctuations drive its abundance farther from the boundary between extant and near-extinct. Therefore the boundary condition at zero abundance is not an absorbing one but rather a permeable boundary where species that hit zero abundance can be supported by small migration and come back into the community soon after their extinction. dPB are able to write dynamical mean field equations for the abundance dynamics and solve these equations numerically to get the steady state abundance distribution. Our case is simpler in this respect since with *evolutionary dynamics* the extinction boundary is strictly absorbing. After guessing a form for the abundance trajectories and enforcing self-consistency, we find that a Markovian *Ansatz* for a purely exponential correlation function of the dynamics yields a self-consistent solution. In our case, the equation satisfied by the abundance distribution reduces to a boundary value problem (Equation I13) which can be solved numerically (Figure 9).

[1] Shinichi Sunagawa, Luis Pedro Coelho, Samuel Chaffron, Jens Roat Kultima, Karine Labadie, Guillem Salazar, Bardya Djahanschiri, Georg Zeller, Daniel R Mende, Adriana Alberti, et al. Structure and function of the global ocean microbiome. *Science*, 348(6237):1261359, 2015.

[2] A framework for human microbiome research. *Nature*, 486 (7402):215–221, 2012.

[3] Noah Fierer. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, 15(10):579–590, 2017.

[4] Nadav Kashtan, Sara E Roggensack, Sébastien Rodrigue, Jessie W Thompson, Steven J Biller, Allison Coe, Huiming Ding, Pekka Marttinen, Rex R Malmstrom, Roman Stocker, et al. Single-cell genomics reveals hundreds of co-existing subpopulations in wild prochlorococcus. *Science*, 344(6182):416–420, 2014.

[5] Michael J Rosen, Michelle Davison, Devaki Bhaya, and Daniel S Fisher. Fine-scale diversity and extensive recombination in a quasisexual bacterial population occupying a broad niche. *Science*, 348(6238):1019–1023, 2015.

[6] R Frank Rosenzweig, RR Sharp, David S Treves, and Julian Adams. Microbial evolution in a simple unstructured environment: genetic differentiation in escherichia coli. *Genetics*, 137(4):903–917, 1994.

[7] Maren L Friesen, Gerda Saxer, Michael Travisano, and Michael Doebeli. Experimental evidence for sympatric ecological diversification due to frequency-dependent competition in escherichia coli. *Evolution*, 58(2):245–260, 2004.

[8] Ram Maharjan, Shona Seeto, Lucinda Notley-McRobb, and Thomas Ferenci. Clonal adaptive radiation in a constant environment. *Science*, 313(5786):514–517, 2006.

[9] Benjamin H Good, Michael J McDonald, Jeffrey E Barrick, Richard E Lenski, and Michael M Desai. The dynamics of molecular evolution over 60,000 generations. *Nature*, 551 (7678):45–50, 2017.

[10] Richard E Lenski. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME journal*, 11(10):2181–2194, 2017.

[11] Steven E Finkel and Roberto Kolter. Evolution of microbial diversity during prolonged starvation. *Proceedings of the National Academy of Sciences*, 96(7):4023–4027, 1999.

[12] J Cesar Ignacio-Espinoza, Nathan A Ahlgren, and Jed A Fuhrman. Long-term stability and red queen-like strain dynamics in marine viruses. *Nature Microbiology*, 5(2):265–271, 2020.

[13] Marcia F Marston, Francis J Pierciey Jr, Alicia Shepard, Gary Gearin, Ji Qi, Chandri Yandava, Stephan C Schuster, Matthew R Henn, and Jennifer BH Martiny. Rapid diversification of coevolving marine synechococcus and a virus. *Proceedings of the National Academy of Sciences*, 109(12): 4544–4549, 2012.

[14] Angus Buckling and Paul B Rainey. Antagonistic coevolution between a bacterium and a bacteriophage. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1494):931–936, 2002.

[15] Leigh Van Valen. A new evolutionary law. *Evolutionary Theory*, 1:1, 1973.

[16] J Arjan GM De Visser and Joachim Krug. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*, 15(7):480–490, 2014.

[17] Daniel S Fisher. Inevitability of red queen evolution driven by organismic complexity and simple feedback via environmental modification. *bioRxiv*, pages 2021–09, 2021.

[18] Atish Agarwala and Daniel S Fisher. Adaptive walks on high-dimensional fitness landscapes and seascapes with distance-dependent statistics. *Theoretical Population Biology*, 130:13–49, 2019.

[19] Guy Bunin. Ecological communities with lotka-volterra dynamics. *Physical Review E*, 95(4):042414, 2017.

[20] Manfred Opper and Sigurd Diederich. Phase transition and 1/f noise in a game dynamical model. *Physical Review Letters*, 69(10):1616, 1992.

[21] Emmy Blumenthal, Jason W Rocks, and Pankaj Mehta. Phase transition to chaos in complex ecosystems with non-reciprocal species-resource interactions. *Physical Review Letters*, 132(12):127401, 2024.

[22] Valentina Ros, Felix Roy, Giulio Biroli, Guy Bunin, and Ari M Turner. Generalized lotka-volterra equations with random, nonreciprocal interactions: The typical number of equilibria. *Physical Review Letters*, 130(25):257401, 2023.

[23] Michael T Pearce, Atish Agarwala, and Daniel S Fisher. Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. *Proceedings of the National Academy of Sciences*, 117(25):14572–14583, 2020.

[24] Chi Xue and Nigel Goldenfeld. Coevolution maintains diversity in the stochastic kill the winner model. *Physical Review Letters*, 119(26):268101, 2017.

[25] Jef Huisman, Anna M Johansson, Eelke O Folmer, and Franz J Weissing. Towards a solution of the plankton paradox: the importance of physiology and life history. *Ecology Letters*, 4(5):408–411, 2001.

[26] Le Yan, Richard A Neher, and Boris I Shraiman. Phylodynamic theory of persistence, extinction and speciation of rapidly adapting pathogens. *Elife*, 8:e44205, 2019.

[27] Robert A Armstrong and Richard McGehee. Competitive exclusion. *The American Naturalist*, 115(2):151–170, 1980.

[28] Noam Shoresh, Matthew Hegreness, and Roy Kishony. Evolution exacerbates the paradox of the plankton. *Proceedings of the National Academy of Sciences*, 105(34):12365–12369, 2008.

[29] Peter Schippers, Antonie M Verschoor, Matthijs Vos, and Wolf M Mooij. Does supersaturated coexistence resolve the paradox of the plankton? *Ecology letters*, 4(5):404–407, 2001.

[30] Michael Doebeli and Iaroslav Ispolatov. Complexity and diversity. *Science*, 328(5977):494–497, 2010.

[31] Mikhail Tikhonov and Remi Monasson. Innovation rather than improvement: A solvable high-dimensional model highlights the limitations of scalar fitness. *Journal of Statistical Physics*, 172:74–104, 2018.

[32] Carlos A Serván and Stefano Allesina. Tractable models of ecological assembly. *Ecology Letters*, 24(5):1029–1037, 2021.

[33] José A Capitán, José A Cuesta, and Jordi Bascompte. Statistical mechanics of ecosystem assembly. *Physical Review Letters*, 103(16):168101, 2009.

[34] Ada Altieri, Felix Roy, Chiara Cammarota, and Giulio Biroli. Properties of equilibria and glassy phases of the random lotka-volterra model with demographic noise. *Physical Review Letters*, 126(25):258301, 2021.

[35] Mikhail Tikhonov and Remi Monasson. Collective phase in resource competition in a highly diverse ecosystem. *Physical Review Letters*, 118(4):048103, 2017.

[36] Wenping Cui, Robert Marsland III, and Pankaj Mehta. Effect of resource dynamics on species packing in diverse ecosystems. *Physical Review Letters*, 125(4):048101, 2020.

[37] Giulio Biroli, Guy Bunin, and Chiara Cammarota. Marginally stable equilibria in critical ecosystems. *New Journal of Physics*, 20(8):083051, 2018.

[38] Robert Marsland III, Wenping Cui, and Pankaj Mehta. The minimum environmental perturbation principle: A new perspective on niche theory. *The American Naturalist*, 196 (3):291–305, 2020.

[39] Juan A Bonachela, Meike T Wortel, and Nils Chr Stenseth. Eco-evolutionary red queen dynamics regulate biodiversity in a metabolite-driven microbial system. *Scientific Reports*, 7(1):17655, 2017.

[40] Richard Law, Paul Marrow, and Ulf Dieckmann. On evolution under asymmetric competition. *Evolutionary Ecology*, 11:485–501, 1997.

[41] Jan Martin Nordbotten and Nils C Stenseth. Asymmetric ecological conditions favor red-queen type of continued evolution over stasis. *Proceedings of the National Academy of Sciences*, 113(7):1847–1852, 2016.

[42] Robert M May. Will a large complex system be stable? *Nature*, 238(5364):413–414, 1972.

[43] Aditya Mahadevan, Michael T Pearce, and Daniel S Fisher. Spatiotemporal ecological chaos enables gradual evolutionary diversification without niches or tradeoffs. *Elife*, 12: e82734, 2023.

[44] Anna Posfai, Thibaud Taillefumier, and Ned S Wingreen. Metabolic trade-offs promote diversity in a model ecosystem. *Physical Review Letters*, 118(2):028103, 2017.

[45] Rodrigo Caetano, Yaroslav Ispolatov, and Michael Doebeli. Evolution of diversity in metabolic strategies. *Elife*, 10: e67764, 2021.

[46] Aisling J Daly, Jan M Baetens, and Bernard De Baets. Ecological diversity: measuring the unmeasurable. *Mathematics*, 6(7):119, 2018.

[47] Frédéric Vrins. Sampling the multivariate standard normal distribution under a weighted sum constraint. *Risks*, 6(3): 64, 2018.

[48] Sigurd Diederich and Manfred Opper. Replicators with random interactions: A solvable model. *Physical Review A*, 39 (8):4333, 1989.

[49] Felix Roy, Giulio Biroli, Guy Bunin, and Chiara Cammarota. Numerical implementation of dynamical mean field theory for disordered systems: Application to the lotka–volterra model of ecosystems. *Journal of Physics A: Mathematical and Theoretical*, 52(48):484001, 2019.

[50] Leticia F Cugliandolo, Jorge Kurchan, Pierre Le Doussal, and Luca Peliti. Glassy behaviour in disordered systems with nonrelaxational dynamics. *Physical Review Letters*, 78(2):350, 1997.

[51] Thibaut Arnoulx de Pirey and Guy Bunin. Many-species ecological fluctuations as a jump process from the brink of extinction. *Physical Review X*, 14(1):011037, 2024.

[52] Madeleine Bonsma-Fisher and Sidhartha Goyal. Dynamics of immune memory and learning in bacterial communities. *Elife*, 12:e81692, 2023.

[53] Jacopo Marchi, Michael Lässig, Aleksandra M Walczak, and Thierry Mora. Antigenic waves of virus-immune co-evolution. *Proceedings of the National Academy of Sciences*, 118(27):e2103398118, 2021.

[54] Akshit Goyal, Leonora S Bittleston, Gabriel E Leventhal, Lu Lu, and Otto X Cordero. Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *Elife*, 11:e74987, 2022.

[55] Takehito Yoshida, Stephen P Ellner, Laura E Jones, Brendan J M Bohannan, Richard E Lenski, and Nelson G Hairston Jr. Cryptic population dynamics: rapid evolution masks trophic interactions. *PLoS Biology*, 5(9):e235, 2007.

[56] Mireia Valles-Colomer, Aitor Blanco-Míguez, Paolo Manghi, Francesco Asnicar, Leonard Dubois, Davide Golzato, Federica Armanini, Fabio Cumbo, Kun D Huang, Serena Manara, et al. The person-to-person transmission landscape of the gut and oral microbiomes. *Nature*, 614 (7946):125–135, 2023.

[57] Wei Zhou, Michelle Spoto, Rachel Hardy, Changhui Guan, Elizabeth Fleming, Peter J Larson, Joseph S Brown, and Julia Oh. Host-specific evolutionary and transmission dynamics shape the functional diversification of staphylococcus epidermidis in human skin. *Cell*, 180(3):454–470, 2020.

[58] Antonio M Martin-Platero, Brian Cleary, Kathryn Kauffman, Sarah P Preheim, Dennis J McGillicuddy, Eric J Alm, and Martin F Polz. High resolution time series reveals cohesive but short-lived communities in coastal plankton. *Nature Communications*, 9(1):266, 2018.

[59] Luigi M Ricciardi and Shunsuke Sato. First-passage-time density and moments of the ornstein-uhlenbeck process. *Journal of Applied Probability*, 25(1):43–57, 1988.